ANALYSIS AND MITIGATION OF OPAMP NONLINEARITIES IN CONTINUOUS-TIME OVERSAMPLED CONVERTERS

A THESIS

submitted by

PRABU SANKAR THIRUGNANAM

for the award of the degree

of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, MADRAS.

August 2010

This thesis is
dedicated to
my parents

Thirugnanam
and
Thenmozhi

THESIS CERTIFICATE

This is to certify that the thesis titled **ANALYSIS AND MITIGATION OF OPAMP NONLINEARITIES IN CONTINUOUS-TIME OVERSAMPLED CONVERTERS**, submitted by **Prabu Sankar T.**, to the Indian Institute of Technology, Madras, for the award of the degree of **Doctor of Philosophy**, is a bona fide record of the research work done by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Y. Shanthi Pavan
Research Guide
Professor
Dept. of Electrical Engineering
IIT Madras, 600 036

Place: Chennai

Date:

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ADC	Analog-to-Digital Converter
DAC	Digital-to-Analog Converter
DSP	Digital Signal Processor
CTDSM	Continuous-Time Delta Sigma Modulator
PSD	Power Spectral Density
SNR	Signal-to-Noise Ratio
SAR	Successive Approximation Register
STF	Signal Transfer Function
NTF	Noise Transfer Function
CIFF	Cascade of Integrators Feedforward
CIFB	Cascade of Integrators Feedback
MSA	Maximum Stable Amplitude
OBG	Out-of-Band Gain
NRZ	Non-Return-to-Zero
RZ	Return-to-Zero
SCR	Switched Capacitor Resistor
DEM	Dynamic Element Matching
SR	Slew-Rate
IBN	In-band Noise

ZOH	Zero-Order-Hold
LSB	Least Significant Bit
SNDR	Signal-to-Noise-plus-Distortion Ratio
DR	Dynamic Range
FOM	Figure of Merit

NOTATIONS

English symbols

f_s	Sampling frequency
T_s	Sampling period
f_b	Signal bandwidth
g_m	Small-signal transconductance
g_3	Third order coefficient of nonlinearity of a transconductor
r_1, r_2	Small-signal resistances of an opamp
c_1, c_2	Small-signal parasitic capacitances of an opamp
V_{ref}	Reference voltage of a DAC
A_{in}	Input signal amplitude
f_{in}	Input signal frequency
HD_3	Third Harmonic Distortion

Greek symbols

Δ	Step size of a quantizer
β	Third order coefficient of nonlinearity in a nonlinear function
σ	Standard deviation of a random variable from its mean
τ	Time constant of an RC network

ACKNOWLEDGEMENTS

At the outset, I would like to express my sincere gratitude to my research advisor Dr. Shanthi Pavan. It was his constant motivations and encouragements that have been instrumental in shaping my research work. I would say I am privileged to not only benefit from his brilliant teaching but also get his invaluable guidance throughout my research life here at IIT, Madras. I would also like to thank him for spending lot of his precious time in helping me bring this thesis in its present form.

I also take this opportunity to thank Prof. Nagendra Krishnapura for imparting significant knowledge in the field of data converters and analog IC design. I am also grateful to him for having many timely technical discussions and for patiently clarifying many of my doubts. I would also like to extend my heartfelt gratitude to Prof. P.V. Ramakrishna (Anna University) at this juncture. His directions and inspirations during my under-graduate years have been significant in beginning my research career.

I am thankful to the members of my Doctoral Committee - Prof. Vinita Vasudevan and Prof. Nagendra Krishnapura of the Electrical Department, Prof. Babu Viswanathan of the Mechanical Department and Dr. Shankar Balachandran from the Computer Science Department, for their comments and suggestions.

I would like to acknowledge Anand Kannan, Sunil Rafeeque and Nitesh Kekre of Texas Instruments for granting access to some high-end test equipments and helping me test my fabricated chips.

I am greatly thankful to all my labmates who have been around me in different stages of my five year stay at IIT, Madras. Some of those lively discussions (technical and non-technical) I had with my senior labmates - Reddy Karthikeyan, Murali, I.Rajesh, Laxmi Nidhi and V.V Srinivas, still remain fresh in my mind. I'm sure they will remain so forever. Equally good were the times spent with my juniors - Deepa, Shankar, Vikas, Hariharan, Amrith, Pradeep, Mrunmay, Lokesh, Eeshan, Saurabh, Sandeep, Radha, Sreelakshmi, Ankesh... (the list goes on). Some of my most enjoyable moments were with this lot, especially Deepa, Vikas and Amrith.

I would also like to thank all the support staff in our department and in the academic section for all the help they extended in administrative matters.

I also take this opportunity to thank my sisters and other family members for their timely moral support. Last but not the least, i would like to mention about my beloved parents, Thirugnanam and Thenmozhi, without whom I wouldn't have been what I am today. It would only be befitting to dedicate this thesis to them for all their care and support they have poured on me, right from day one I stepped on this planet.

(Prabu Sankar T.)

ABSTRACT

KEYWORDS: Continuous-time; analog-to-digital; oversampling; delta-sigma; opamp nonlinearity; harmonic distortion; assisted opamp; audio

This research addresses the need to understand the effects of opamp nonlinearity on the performance of a class of oversampled analog-to-digital converters called continuous-time delta-sigma modulators. Nonlinearity in the opamps used to build continuous-time filters are one of the crucial factors determining the power consumption of these modulator. In modulators employing multi-bit quantizers, opamp nonlinearity is found to degrade the modulator performance by increasing the in-band noise. Analytical relations quantifying the increase in the in-band noise spectral density are developed for various integrator architectures and different opamp topologies. The analysis is also extended to quantify the harmonic distortions observed in modulators with single-bit quantizers. Macromodel simulation results are used to prove the accuracy of the derived analytical results.

A power efficient technique that can mitigate the effects of opamp nonlinearity in continuous-time delta-sigma modulators is also proposed. Termed as the 'assisted opamp technique', it can be used to achieve low power, low distortion operation in single-bit modulators. The efficacy of the technique is demonstrated through two single-bit audio modulators - one using a Non-Return-to-Zero (NRZ) feedback DAC and the other using a Switched Capacitor Resistor (SCR) feedback DAC. With the help of the

proposed technique, the modulator with an NRZ DAC is found to achieve a dynamic range of 92.5 dB in a 24 kHz bandwidth, with a power dissipation of 110 μW from a 1.8 V supply. The second design employing an SCR DAC, achieves a dynamic range of 91.5 dB and dissipates 122 μW of power. The performance thus achieved with these modulators is comparable to that of the best multi-bit modulators reported in the literature.

CHAPTER 1

INTRODUCTION

Advancements in the field of integrated circuits over the last few decades has made electronics an indispensable part of human life. With extensive computational capabilities, robustness, repeatability and long-term stability, digital signal processing techniques are becoming increasingly preferred over traditional analog processing. However, the world surrounding us is analog in nature and therefore, there is a need to interface the real world with digital processors. As depicted in Fig. 1.1, Analog-to-Digital Converters (ADCs) and Digital-to-Analog Converters (DACs) form a vital part of the interface electronics that bridge the analog and digital world.

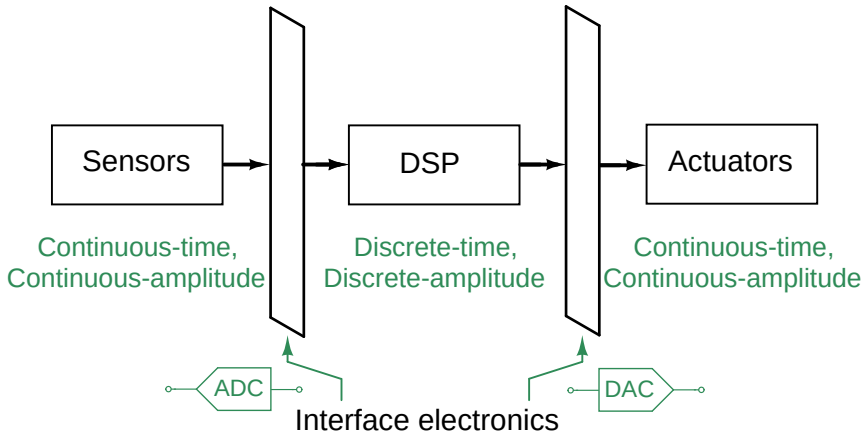


Figure 1.1: Functional view of the role played by ADC and DAC

As the name suggests, the function of ADCs is to convert the continuous-time analog signals to digital signals that can be suitably processed by a Digital Signal Processor (DSP). DACs perform the inverse operation of converting the processed output from the

DSP back to the continuous-time domain. The focus of this thesis is a popular 'oversampled' ADC architecture called the *Continuous-Time Delta Sigma Modulator (CTDSM)*. This chapter is aimed at providing some of the basic concepts that can aid better understanding of the ideas and models presented in this thesis and has been organized as follows. Beginning with the fundamentals of analog-to digital conversion, the discussion is carried over to popular ADC architectures that are currently being used. Subsequently, the reader is introduced to delta sigma modulators, of which continuous-time modulators form a sub-class. Operational details and various implementation issues involving these modulators are then described, eventually bringing about the motivation behind the problem addressed in this thesis. The chapter is concluded by describing the organization of the thesis and contributions of this research work.

1.1 Fundamentals of ADC - Sampling and Quantization

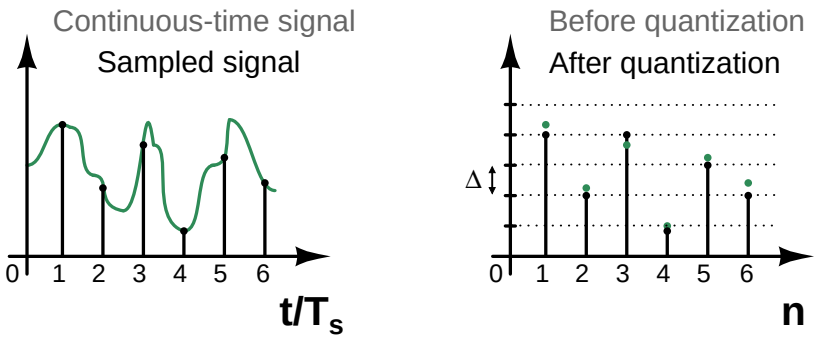


Figure 1.2: Illustration of sampling and quantization operations occurring in an ADC

The conversion from the analog to digital domain achieved in an ADC involves two fundamental operations - *sampling* and *quantization*. As illustrated in Fig. 1.2, sampling involves the conversion of the continuous-time analog signal to a discrete-time signal, sampled every T_s secs. The next process involves quantization of these samples to have discrete amplitude levels (in steps of Δ as shown in Fig. 1.2). The discrete-levels are

then appropriately mapped to the digital format to be able to be processed by a DSP.

In a nutshell, ADCs can be perceived as performing discretization in both time and amplitude domains. Such discretizations can be expected to lead to potential loss of information contained in the analog signal. Information loss during the sampling process is averted by adhering to Shannon's sampling theorem according to which '*any analog signal with a bandwidth of f_b can be completely reconstructed from its samples, if the sampling rate (f_s) is more than $2f_b$*'. The critical sampling rate of '$2f_b$' is termed as the Nyquist-rate.

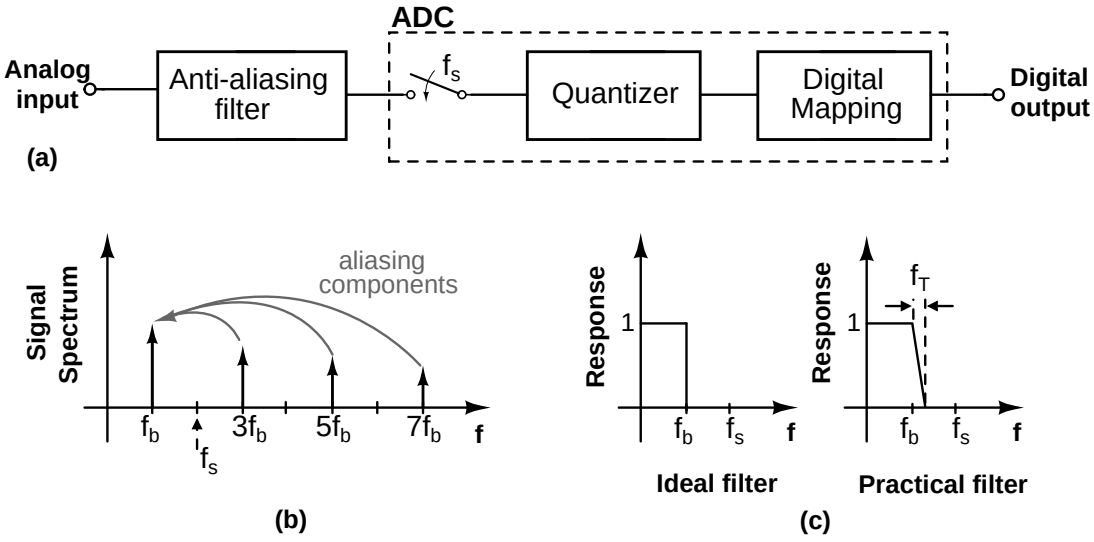


Figure 1.3: (a) Model of an ADC along with anti-aliasing filter (b) Illustration of aliasing into the signal band from components around multiples of $\pm f_s$ (c) Response of ideal and practical anti-aliasing filters

Since not all signals are band-limited in nature, a low pass filter (called the 'anti-aliasing' filter) is always required to precede the sampling operation of an ADC, as shown in Fig. 1.3(a). Aliasing of frequency components around multiples of $\pm f_s$ into the signal band (as illustrated in Fig. 1.3(b)), can be prevented by the anti-aliasing filter. Though an ideal brick-wall response is desired out of an anti-aliasing filter, practical realizability issues force the use of filters with finite transition bands (f_T) as shown in

Fig. 1.3(c). The order and the transition band requirements on the filter are determined by the nature (or the frequency content) of the continuous-time signal and the sampling rate used. Naturally, such specifications are the most severe when signals are sampled at the Nyquist-rate.

While discretization in the time-domain can be shown to be lossless, the same cannot be said about the subsequent operation of quantization. Discretization in the amplitude domain inherently introduces 'permanent' errors called the quantization error or quantization noise, thereby resulting in a lossy conversion. Such errors can be reduced by increasing the number of discrete-levels a quantizer can resolve into, or in other words, by reducing the step size (Δ) that separates each of these discrete-levels.

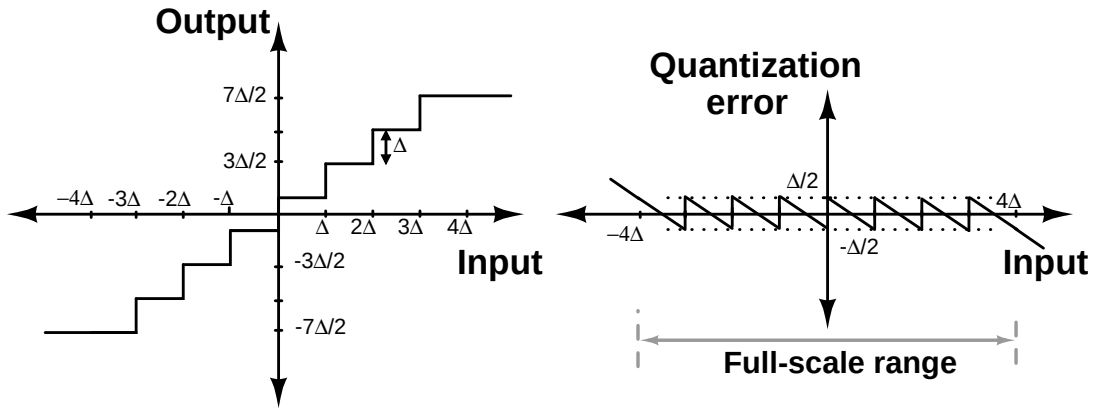


Figure 1.4: Typical input-output characteristic of a 8-level quantizer along with the quantization error profile

Illustrating these two design parameters of a quantizer is Fig. 1.4, where the input-output characteristic of a 8-level quantizer and the corresponding quantization error are shown. The quantization error can be seen to be bounded between $-\Delta/2$ and $\Delta/2$ for inputs in the range [-4Δ, 4Δ]. This input range, beyond which the quantization error can be seen to grow without bounds is called the full-scale range of the quantizer. The quantizer is said to be saturated or over-loaded for inputs above this range.

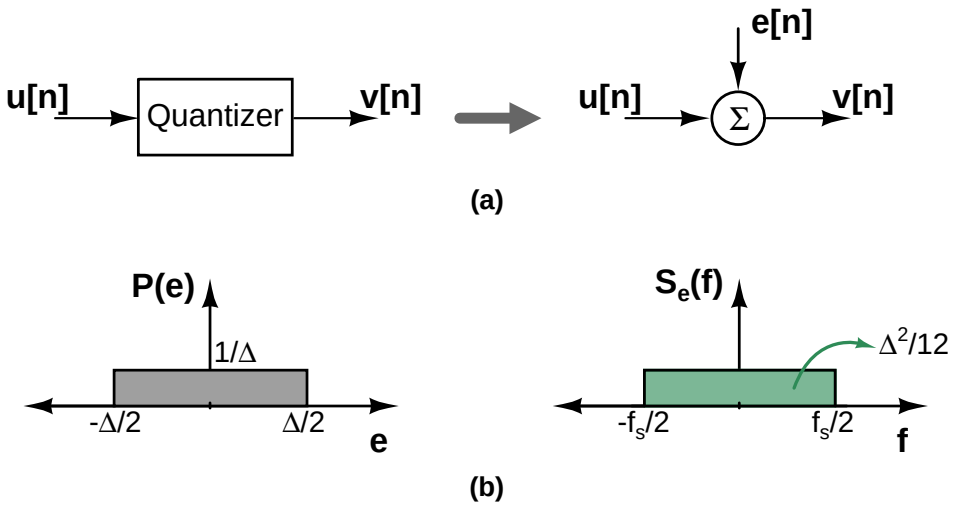


Figure 1.5: (a) Linear representation of a quantizer (b) Probability density function, $P(e)$ and Power Spectral Density, $S_e(f)$ of the quantization noise

Though quantization is a nonlinear operation, dynamic performance of quantizers (and ADCs) has been traditionally analyzed by adopting a linearized, stochastic model as shown in Fig. 1.5(a). The quantizer is perceived as an independent additive source that introduces the quantization noise sequence, $e[n]$. As shown in Fig. 1.5(b), $e[n]$ is assumed to be uniformly distributed (Bennett (1948)) in the range [-Δ/2, Δ/2], with a white Power Spectral Density (PSD), under the following conditions[1]

- The quantizer input is within the full scale range
- The number of quantizer levels is large and the step size, Δ is very small
- The input is random in nature

Under such assumptions, the quantization noise can be shown to have a variance of $\Delta^2/12$ and a PSD given by $\Delta^2/(12f_s)$. Though the white noise approximation becomes invalid for sinusoidal inputs (since the quantization noise becomes periodic), the quantization noise power is found to be close enough to $\Delta^2/12$. Therefore, the dynamic performance of an ADC is measured by the Signal-to-Noise Ratio (SNR) observed for a sinusoidal input, assuming a variance of $\Delta^2/12$ for the quantization noise. For a

[1]While these conditions serve to provide a general outlook on the cases when the above assumption is valid, more specific conditions on the characteristics of the input signal can be found in Widrow (1956) and Sripad and Snyder (1977).

full-scale input signal, the resulting peak SNR can be shown to be

$$SNR = (6.02N + 1.76)\,\mathrm{dB} \tag{1.1}$$

where N denotes the number of bits used to represent the discrete levels of the quantizer[2]. The above equation is often used in the reverse sense to define the resolution of an ADC, as the value of N determined by the peak SNR achieved with the ADC. The resolution increases by one bit for every 6 dB increase in the peak SNR.

1.2 ADC architectures

Various ADC architectures have been proposed to accomplish the conversion from analog to digital domain based on the principles discussed above. Based on the sampling rate, they can be broadly classified into Nyquist-rate and oversampled converters.

A. Nyquist-rate Converters : Nyquist-rate converters sample the input signal at twice the signal bandwidth ($2f_b$). Some of the popular architectures are (Johns and Martin (2009))

- Flash ADCs
- Successive Approximation Register (SAR) ADCs
- Pipeline ADCs

The building block behind all these architectures is the *comparator*, which decides whether the input signal is above or below a threshold and accordingly generates a binary '1' or '0' as the output. Each architecture employs the comparator in an unique manner to facilitate the required analog to digital conversion.

[2]In effect, an ADC with 2^N discrete levels of quantization has N bits in its digital output.

For instance, a N-bit flash ADC employs 2^N-1 comparators that act in a concurrent fashion, to achieve the conversion. Fig. 1.6(a) shows a typical 2-bit flash ADC built with 3 comparators, with each of the thresholds tapped from the resistive ladder shown. The comparators together provide a thermometer coded output[3] which is then decoded to obtain a 2-bit binary output.

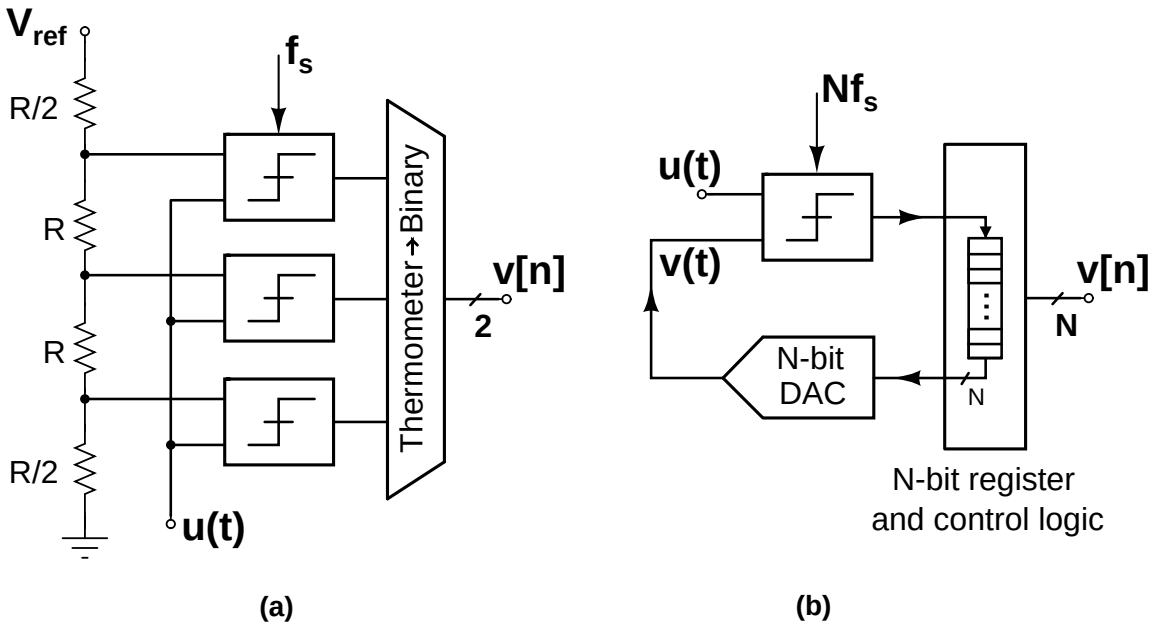


Figure 1.6: (a) 4-bit Flash ADC (b) Functional block diagram of a SAR ADC

Due to the exponential increase in the number of comparators with increasing resolution (N), flash ADCs are not used to achieve resolutions higher than 8 bits. For some applications requiring higher resolutions, SAR ADCs can be used. As shown in Fig. 1.6(b), SAR ADCs employ a single comparator in an iterative fashion. They use the *binary search* algorithm to systematically reduce the difference between the input signal ($u(t)$) and its estimate given by the fed back DAC signal ($v(t)$). Such a digitization process requires N clock-cycles to generate a N-bit digital output. This clearly shows that speed is traded off to have reduced complexity (area), when adopting SAR ADCs over flash

[3]Similar to the dipping of mercury level of a thermometer, the comparators whose thresholds are lower than the input signal, generate a logic high (or '1') at their outputs.

ADCs. For applications that require higher data rates than supported by SAR ADCs,

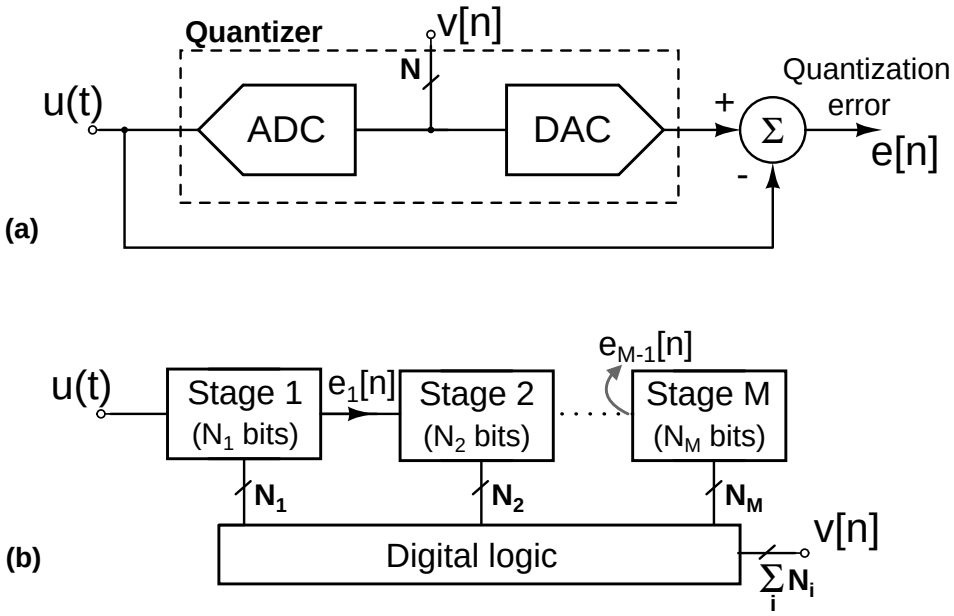


Figure 1.7: (a) Individual stage of a pipeline ADC (b) Functional block diagram of a pipeline ADC having M stages

Pipeline ADCs can be used. As the name suggests, this architecture utilizes the principle of pipelining to engage M stages of conversion, in a concurrent fashion. As depicted in Fig. 1.7(a), each stage employs an ADC (typically a flash) and a DAC to generate a N-bit digitized output along with the quantization error, $e[n]$. As shown in Fig. 1.7(b), the quantization error generated by each stage ($e_i[n]$) is then passed on to the subsequent stage, for digitization in the 'next' clock cycle. The digitization of a given input sample is completed after M clock cycles (stages), when a digital block appropriately combines the output of each stage to generate the desired digital output. Except for the initial latency period[4] of typically M clock cycles, very high conversion rates can be achieved with this architecture, thanks to the parallel conversion operations occurring in each of the M stages. It can be seen that circuit complexity has been traded off to achieve higher speeds with a pipeline architecture.

[4]This refers to the number of clock cycles needed before the first digital output can be obtained corresponding to the first input sample.

B. Oversampled converters : While Nyquist-rate converters can be said to be strictly following the sampling theorem, analog signals can be digitized at a much higher sampling rate as is done in *oversampled converters*. Oversampling results in several benefits such as

- The requirements on the anti-aliasing filter are relaxed due to oversampling, as they can have less-steeper roll-off characteristics with $f_s/2 \gg f_b$.
- More importantly, the quantization noise (with variance $\Delta^2/12$) is now spread across a wider bandwidth. This effectively reduces the quantization noise power in the desired signal bandwidth[5] (f_b), as shown in Fig. 1.8(a).

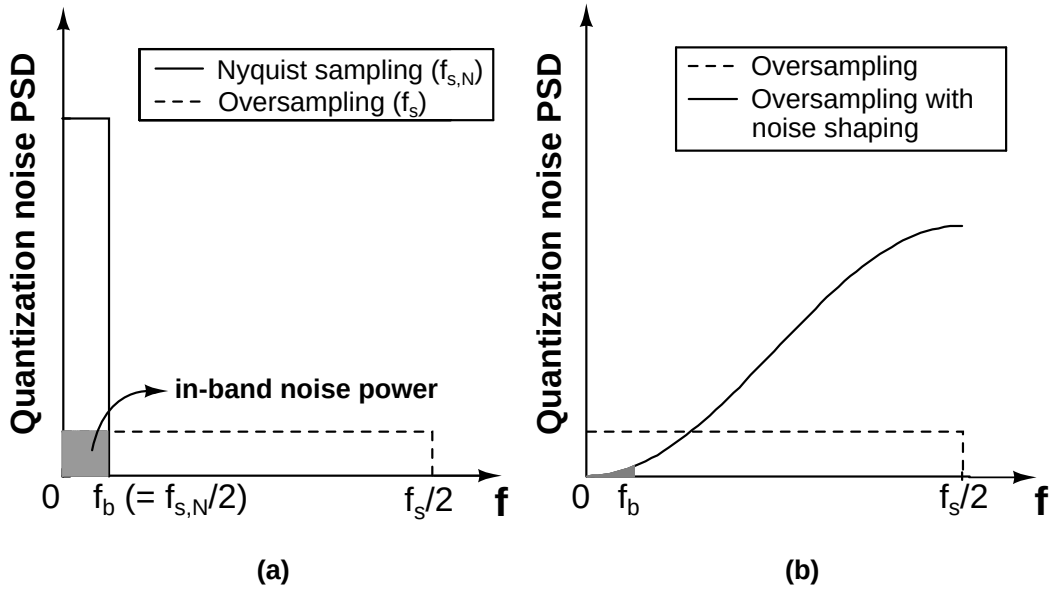


Figure 1.8: (a) Comparison of quantization noise PSD (single-sided) with Nyquist-rate sampling and oversampling (b) Noise shaping achieved in delta-sigma modulators

In essence, oversampling is an useful technique that can increase the effective resolution of an ADC. The output of an oversampled converter can be filtered to remove the noise outside the signal band of interest and then re-sampled at the Nyquist rate to achieve the desired improvement in resolution. Defining a factor called Over-Sampling Ratio (OSR = $f_s/(2f_b)$), increasing the OSR by a factor of 2 can be seen to result in a 3 dB reduction in the in-band quantization noise power. However, observe that for

[5]The desired signal band will henceforth be referred to as the in-band region

every half a bit of improvement in the effective resolution, the sampling rate has to be increased in an exponential manner. Thus, it should be mentioned that additional techniques have to be employed along with oversampling, to obtain significant improvements in the resolution of a converter. Delta-Sigma modulators are one such sub-class of oversampled converters that can achieve very high resolutions, by employing the technique of 'noise shaping' (as depicted in Fig. 1.8(b)). While pipeline architectures can be said to increase the speed of conversions, delta-sigma modulators are seen as suitable choices to achieve high resolutions, that are hard to achieve with a SAR ADC. The following sections have been organized to provide a more detailed overview about this architecture.

1.3 Basics of Delta-Sigma modulation

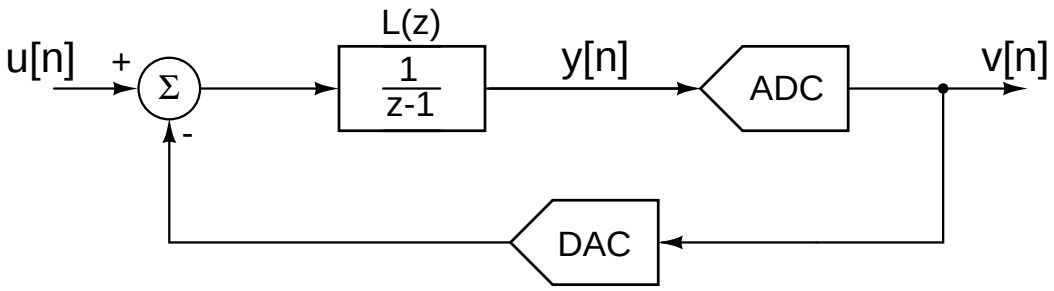


Figure 1.9: Functional block diagram of a first order Delta-Sigma modulator

Fundamentals of delta-sigma modulators can be best understood by considering the first-order modulator as shown in Fig. 1.9. The modulator comprises of an accumulator[6], described by $L(z) = 1/(z-1)$ and an internal, low resolution quantizer (ADC-DAC) that together form a negative feedback system. Analysis of such closed loop systems is in principle complicated due to the presence of the nonlinear quantizer. Nevertheless, a qualitative understanding of the modulator operation can be achieved by

[6]typically implemented using switched-capacitor integrators (Norsworthy *et al.* (1997)).

assuming the traditional linearized model for the quantizer, as shown in Fig. 1.10.

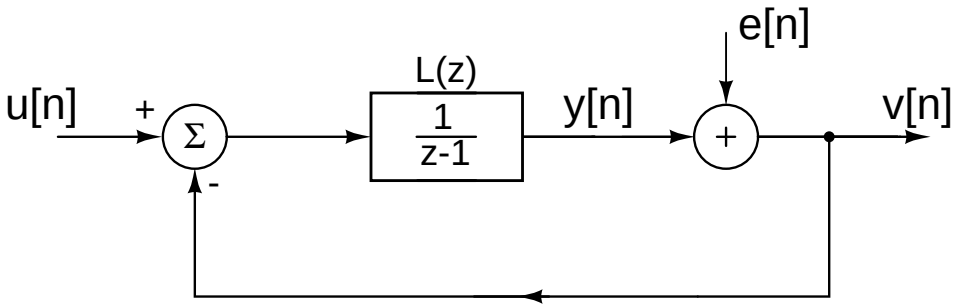


Figure 1.10: Linearized model of a first order delta-sigma modulator

As before, the quantization noise $e[n]$ is assumed to be an independent, uniformly distributed, white random process[7]. Thus, the modulator can be perceived as a system with two inputs, $u[n]$ and $e[n]$ and its output can be expressed in the frequency domain as

$$V(z) = \frac{L(z)}{1+L(z)}U(z) + \frac{1}{1+L(z)}E(z) \tag{1.2}$$

By defining two independent transfer functions namely, the Signal Transfer Function (STF) and the Noise Transfer Function (NTF), we have

$$STF(z) = \frac{L(z)}{1+L(z)} = z^{-1} \tag{1.3}$$

$$NTF(z) = \frac{1}{1+L(z)} = 1 - z^{-1} \tag{1.4}$$

While $|STF|$ is one for in-band signals ($f_b \ll f_s$), the NTF can be seen to be high pass in nature with a transmission zero at DC. Such a shaping of the quantization noise is achieved because of the negative feedback around the quantizer. In conjunction with oversampling, this specifically helps in achieving very high effective resolution. The

[7]Modulators with multi-bit quantizers largely satisfy the conditions needed for such an assumption (Gray (1990))

high resolution thus obtained can be visualized as the modulator's ability to closely track the input signal, thanks to the high gain provided by the accumulator in the in-band region. Illustrating this close tracking is Fig. 1.11, where the input-output waveforms of a first order modulator (with a 4-bit internal quantizer) have been plotted. Due to the oversampled operation, delta-sigma modulators trade off speed to achieve very high resolution, similar to a SAR ADC. However, it can be seen that the improvement achieved with delta-sigma modulators is manifold, because of the noise shaping achieved using negative feedback.

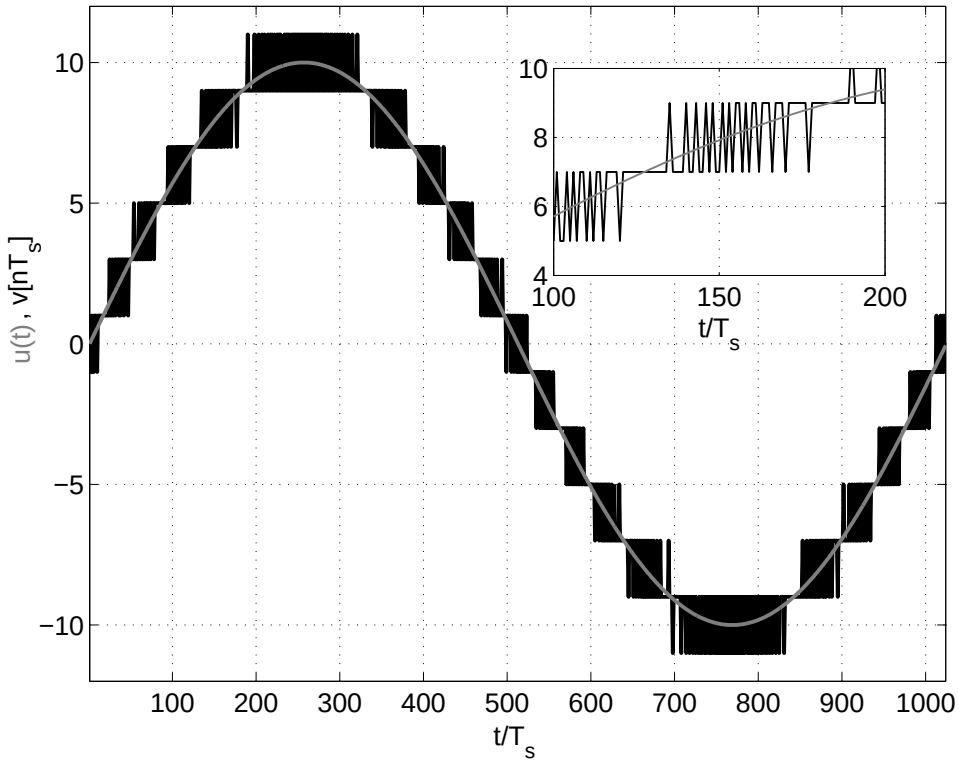


Figure 1.11: Input and output waveforms of a 4-bit modulator employing an accumulator as the loop filter. The inset is a zoomed in version showing how well the output signal tracks the input. The quantizer step size is 2.

Though the above discussion explained delta-sigma concepts with a first order modulator, better noise shaping and higher resolutions can be achieved by using higher order modulators. In general, higher order modulators can be represented as shown

in Fig. 1.12, where the loop filter has two inputs ($u[n]$ and $-v[n]$) with the respective transfer functions being $L_0(z)$ and $L_1(z)$.

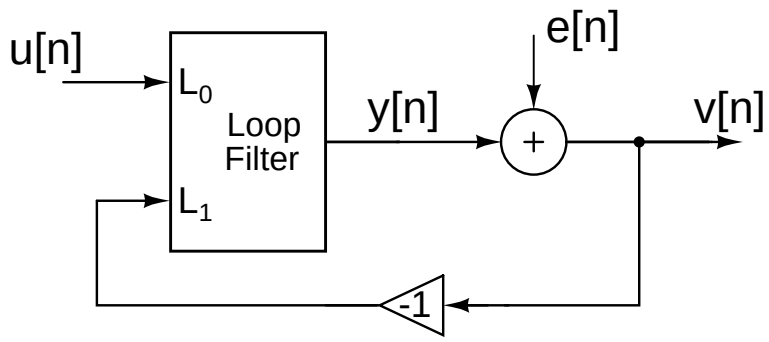


Figure 1.12: Linearized model for higher order delta-sigma modulators

The corresponding STF and NTF are given by

$$STF(z) = \frac{L_0(z)}{1 + L_1(z)} \tag{1.5}$$

$$NTF(z) = \frac{1}{1 + L_1(z)} \tag{1.6}$$

In practice, a suitable choice for the noise and signal transfer functions is the starting point for the design of delta-sigma modulators. Given the desired NTF and STF, the required loop filter transfer functions can be determined as

$$L_0(z) = \frac{STF(z)}{NTF(z)} \tag{1.7}$$

$$L_1(z) = \frac{NTF(z) - 1}{NTF(z)} \tag{1.8}$$

From the above equations, it can be seen that the zeros of the NTF are the poles of both L_0 and L_1. Therefore, circuitry is generally shared when realizing the two transfer functions of the loop filter. Two such loop filter architectures that are commonly used are as shown in Fig. 1.13(a) and (b). They are termed as the

a.) Cascade of Integrators Feedforward (CIFF) architecture[8]

b.) Cascade of Integrators Feedback (CIFB) architecture

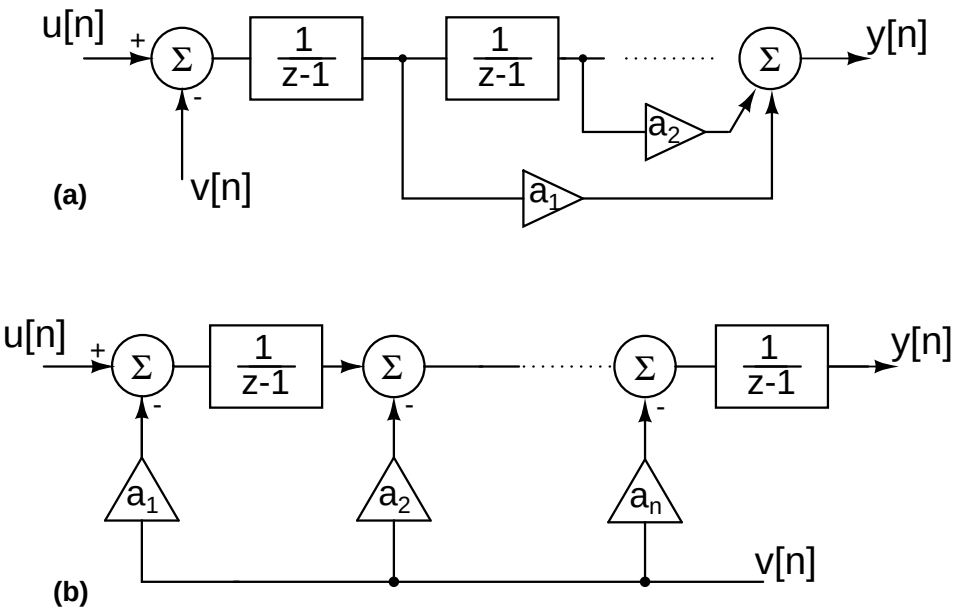


Figure 1.13: (a) Cascade of Integrators Feedforward (CIFF) architecture (b) Cascade of Integrators Feedback (CIFB) architecture

While each of the loop filter architectures has its own advantages and disadvantages (as elaborated in Schreier and Temes (2005*a*)), it should be mentioned that the fundamental operation of the modulator is solely determined by the NTF, STF and the number of the levels in the quantizer, regardless of the loop filter architecture. In general, in an L^{th} order modulator, the magnitude response of the NTF in the in-band region will be of the form $|1 - e^{-j\omega}|^L$. With such an NTF, the in-band 'shaped' quantization noise power can be shown to be, (Norsworthy *et al.* (1997))

$$\sigma_q^2 = \frac{\Delta^2}{12}\left[\frac{\pi^{2L}}{2L+1}\left(\frac{1}{OSR}\right)^{(2L+1)}\right] \tag{1.9}$$

where Δ denotes the step size of the internal quantizer. From the above equation, it can be observed that the in-band noise decreases by (6L+3) dB for every doubling of

[8]Notice that with a CIFF architecture, $L_0(z) = L_1(z)$.

the OSR. This is equivalent to an improvement of L+0.5 bits in the effective resolution. With a first order modulator having a 4-bit internal quantizer, an estimate indicates that the OSR needs to be 256 to be able to achieve 16-bits of effective resolution[9]. The required OSR can be lowered to as low as 32 with a second order modulator, which clearly indicates the significant improvement possible with higher-order architectures.

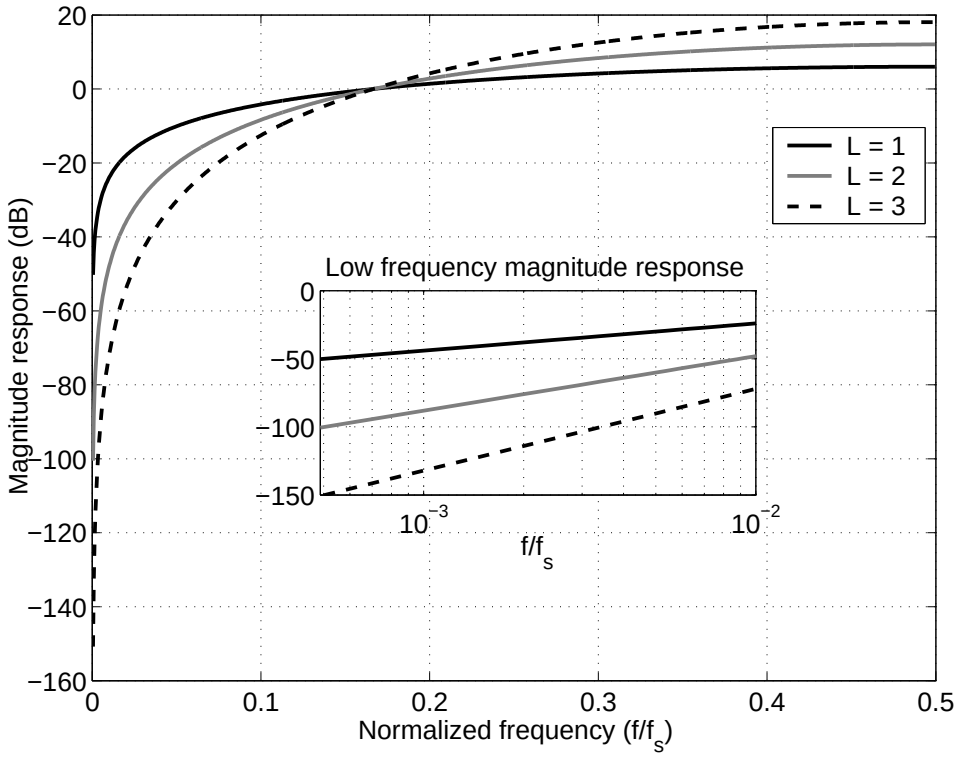


Figure 1.14: Variation in the magnitude response of an all zero NTF ($(1-z^{-1})^L$) for different orders. The inset compares the low frequency responses.

However, the reader must be informed of the potential signal dependent instability that could be observed with higher order modulators. Such stability considerations can be easily understood by considering an all-zero NTF of the form $(1-z^{-1})^L$. As shown in Fig. 1.14, when the order is increased, the in-band noise shaping improves at the cost of an increased out-of-band response for the NTF. In particular, at $\omega = \pi$, it can be seen that $|NTF(e^{j\omega})|$ is 2^{L} for an L^{th} order modulator. This can lead to stability problems

[9]When only oversampling is used, the required OSR should be as high as 2^{24}.

in the modulator operation due to the following reason. The input to the quantizer can be seen to be

$$Y(z) = V(z) - E(z) \tag{1.10}$$

$$= STF(z)U(z) - (NTF(z) - 1)E(z) \tag{1.11}$$

Since the quantizer input consists of the input signal plus the 'shaped' quantization noise, there exists a Maximum Stable Amplitude (MSA)[10] for the input signal, beyond which the quantizer gets saturated. Being a feedback system, the over-load condition of the quantizer only gets worsened further, ultimately forcing the modulator into an 'unstable' region. Physically, for inputs beyond the MSA, the modulator can be thought as having lost its ability to track the input signal. As a result, it is observed to exhibit an unstable behavior, that is marked by some repeatedly failed attempts to track the input signal.

Though discussed with an all-zero NTF, the enhancement of $|NTF(e^{j\omega})|$ at higher frequencies and the subsequent reduction in the MSA can in general be observed with higher order modulators. With these modulators, though the in-band quantization noise can be as expected for smaller inputs, for higher inputs (beyond the MSA), the in-band noise can increase due to instability. This leads to a reduction in the maximum achievable SNR, clearly indicating the potential drawback with higher-order modulators. Another implication of eq. 1.11, worth noting, is the possible reduction in the MSA with reducing number of quantizer levels. The enhancement of the quantization noise, $e[n]$ can be attributed to such a reduction in these cases.

[10]Understandably, the MSA will be smaller than the full-scale range of the quantizer.

While the above discussion suffices to provide an intuitive explanation to the potential issue of stability in delta-sigma modulators, the reader should bear in mind that the analysis was carried out with a linear model for the quantizer. Accurate predictions of the MSA cannot be achieved with such a linear analysis and as often is the case, unforeseen instability conditions can be encountered in practice, at smaller input signal amplitudes. This is especially true when the number of quantizer levels is small (for instance, a single-bit quantizer), since the linear approximation fails in such cases.

Several works have therefore focused on the issue of stability in delta-sigma modulators (as discussed in Schreier and Temes (2005*a*)), in an attempt to demystify the concepts involved in designing stable higher-order modulators. For the benefit of designers, certain thumb-rules have been proposed to successfully utilize the improvement in resolution provided by the higher order modulators. The rules are based on a property of the NTF called the Out-of-Band Gain (OBG), which signifies the maximum gain achieved in the out-of-band region of the NTF. For ensuring stability in modulators with a single-bit quantizer, the OBG is to be restricted to 1.5 (known as Lee's rule -Chao *et al.* (1990)). This condition is relaxed for multi-bit modulators, wherein OBGs typically used are in the range of 2-3. The choice of OBG in such cases depends on various factors including the trade-off between better noise shaping and reduced MSA.

The above conditions on OBG can be satisfied in higher order modulators by appropriately placing the poles of the NTF[11]. Suitable (high pass) filter responses that can be used to meet such requirements are the Butterworth and inverse-Chebyshev type responses. Both these responses exhibit a maximally flat out-of-band region, with the

[11]The poles are realized using the coefficients a_1, a_2, $a_3 \cdots$, with both CIFF and CIFB loop filters shown in Fig. 1.13.

OBG given by the gain of the responses at $\omega = \pi$. The location of the complex transmission zeros in an inverse-Chebyshev response can be optimized to result in better noise shaping and thereby minimize the in-band quantization noise power. Standard routines are available to synthesize the NTF (Schreier's toolbox in MATLAB -Schreier and Temes (2005*a*)), given the loop filter's order, OSR, OBG and the presence/absence of complex in-band zeros. Fig. 1.15 shows the magnitude response of a typical third order NTF realized for an OSR of 128, with complex zeros and an OBG of 1.5.

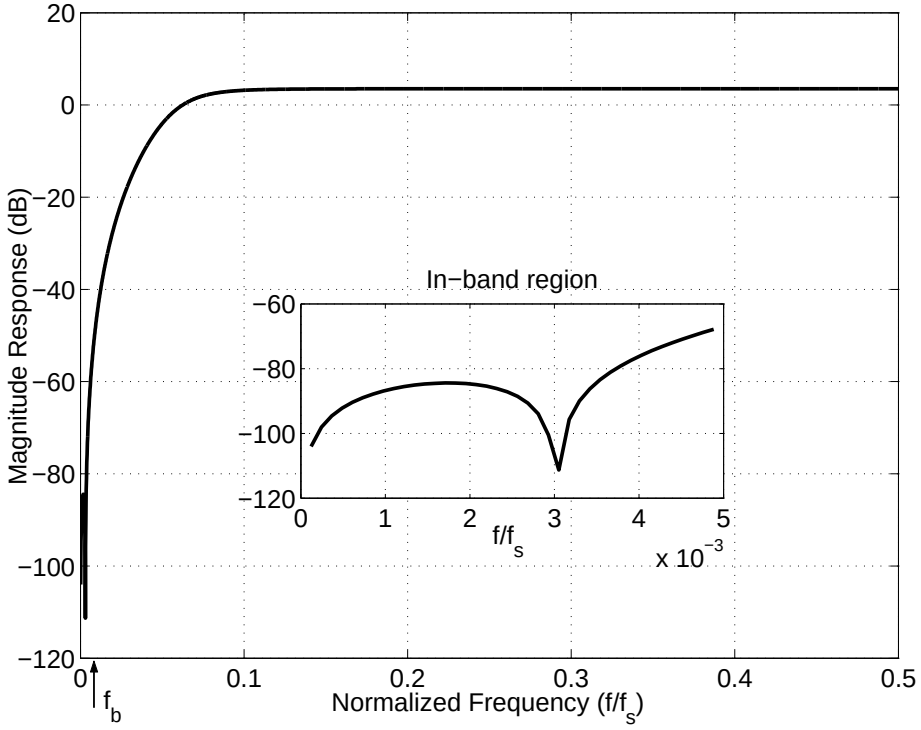


Figure 1.15: Magnitude response of a typical third order NTF with an OBG of 1.5. The inset shows the presence of in-band complex zeros.

The NTF thus obtained can be realized through an appropriate loop filter built using conventional switched-capacitor techniques (Norsworthy *et al.* (1997)). Alternatively, the loop filter can also be implemented in continuous-time. This leads us to a separate class of modulators called Continuous-Time Delta Sigma Modulators (CTDSMs), as shown in Fig. 1.16. With CTDSMs being the focus of this thesis, the rest of the chapter is dedicated to discussions involving their operation, advantages and disadvantages.

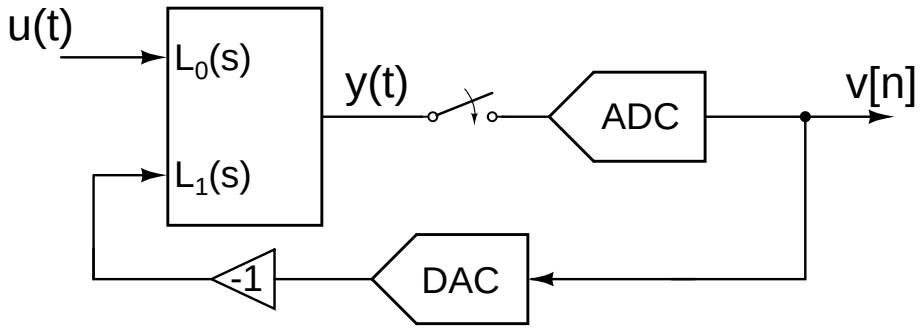


Figure 1.16: Functional block diagram of a Continuous-Time Delta-Sigma Modulator

1.4 Continuous-Time Delta-Sigma Modulators

Historically, delta-sigma modulators were first reported by Inose *et al.* (1962) with a continuous-time loop filter. However, with the advent of switched capacitor circuits in the 1980s, discrete-time implementations were preferred over their continuous-time counterparts for the following reasons

- Switched-capacitor integrators can provide accurate gains, as they are purely determined by capacitor ratios. This is in contrast to the process dependent time constants in continuous-time integrators. As will be explained later, this is observed to lead to stability issues in CTDSMs.
- From a design perspective, since the loop filter in discrete-time is independent of the clock rate, a discrete-time modulator can be operated at different clock frequencies. However, similar to the sensitivity to time-constant variations, a given CTDSM design is only functional over a narrow range of sampling rates.

1.4.1 Advent of continuous-time delta-sigma modulators and their advantages

Attention to continuous-time implementations has gradually increased in the recent past, with the urge to increase the speeds at which delta-sigma modulators work. This is due to the reduced speed requirements on the active circuits of a continuous-time loop filter when compared to that of a switched-capacitor (SC) filter. With an SC integrator, the speed is limited by the settling error which grows exponentially as f_u/f_s (Schreier and Temes (2005*a*)), where f_u is the unity gain frequency of the opamp. In comparison,

the dependency on the f_u/f_s ratio is found to be lesser in continuous-time integrators, thus enabling CTDSMs to operate faster than their discrete-time counterparts. Such a renewed interest has actually thrown light on many other significant advantages offered by CTDSMs. Such advantages seem to have missed the attention of designers, when discrete-time modulators were being preferred. For instance, due to the relaxed speed requirements, CTDSMs consume lesser power than their discrete-time counterparts, to achieve the desired resolution in a given signal bandwidth (f_b). In addition, by virtue of the position where the input is sampled, CTDSMs offer other important benefits like

- Sampling errors at the internal ADC input are shaped away by the loop filter's gain at low frequencies, similar to the quantization noise (van der Zwan and Dijkmans (1996)).
- The loop filter serves as an inherent 'anti-aliasing filter', attenuating any potential aliasing frequencies before they get sampled at the input of the ADC (Candy (1985)).
- The input impedance offered by the modulator is resistive, thus making them easier to drive.

It can be seen that accurate sample and hold (S/H) circuits, buffers and anti-aliasing filters are not required when implementing CTDSMs. Realizing the various system-level and architecture-level advantages accrued by adopting continuous-time modulators, CTDSMs are now being increasingly preferred over discrete-time modulators. Thanks to some extensive research carried out over the recent years, several ways have been devised to tackle various issues associated with their design. This has enabled successful implementations of high-performance CTDSMs, as can be seen from some of the designs reported in the recent past (Mitteregger *et al.* (2006), Ouzounov *et al.* (2007), Reddy and Pavan (2008), Pavan *et al.* (2008)). Before discussing some of the important research works concerning CTDSMs, let us briefly look into the typical design methodology adopted for CTDSMs.

1.4.2 Design methodology of continuous-time modulators

Continuous-time modulators differ from their discrete-time counterparts in the way the loop filter is implemented. However, isolating the loop around the ADC, an equivalence can be found between the continuous-time and the discrete-time system as shown in Fig. 1.17. Simply put, for a given sequence $v[n]$, if the two systems produce the same output sequence $y[n]$, they are said to be equivalent.

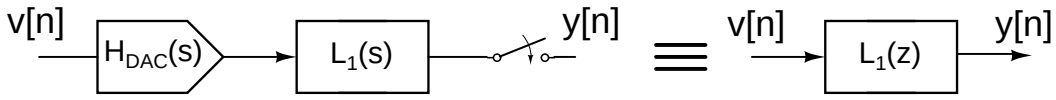


Figure 1.17: Equivalence between continuous-time and discrete-time loop filter ($H_{DAC}(s)$ denotes the Laplace transform of the DAC pulse shape).

Such an equivalence between the two systems can be obtained using the *Impulse-invariant transformation* (Gardner (1986)). With this technique, a stable discrete-time loop filter (with poles inside the unit circle) can be transformed to a stable continuous-time equivalent and vice-versa[12]. Taking note of the possibility of such an equivalence, conventional design strategies for CTDSMs have been primarily based on equivalent discrete-time modulators. This is done to take advantage of the extensive stability analysis carried out for their discrete-time counterparts.

In essence, the first task performed while designing a CTDSM is the *prototyping of the Noise Transfer Function (NTF)*, using the widely popular Schreier's toolbox in MATLAB (as done for a discrete-time design). The discrete-time loop filter, $L_1(z)$ that can realize the NTF is then determined using eq. 1.8. The next step is to apply the Impulse-invariant transformation to determine the equivalent continuous-time filter, $L_1(s)$ for a given shape of the DAC pulse. The equivalence shown in Fig. 1.17 can be

[12]The DAC has been explicitly included in the continuous-time case, as the DAC pulse shape can affect the equivalence.

mathematically expressed as

$$\mathcal{Z}^{-1}\left\{L_1(z)\right\} \;=\; \mathcal{L}^{-1}\left\{H_{DAC}(s)L_1(s)\right\}|_{t=nT_s} \tag{1.12}$$

where $H_{DAC}(s)$ denotes the Laplace transform of the DAC pulse shape.

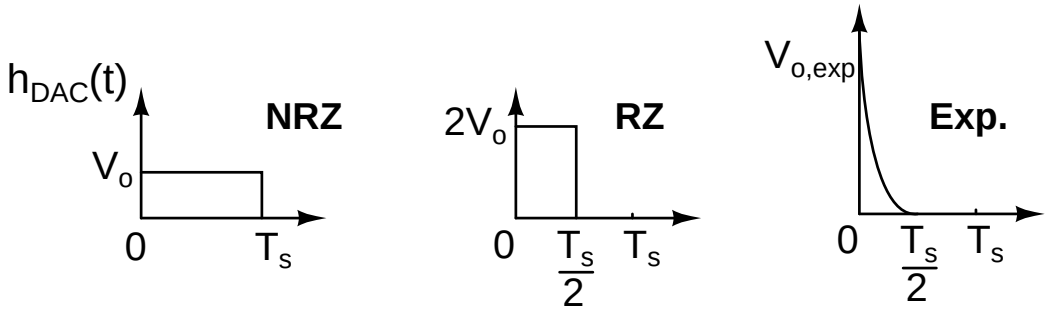


Figure 1.18: Commonly used DAC pulse shapes

Some typical DAC pulse shapes used in practice are shown in Fig. 1.18. They are called Non-Return-to-Zero (NRZ), Return-to-Zero (RZ) and Exponentially decaying (Exp.) pulse shapes. $H_{DAC}(s)$ for the respective pulse shapes are given by

$$H_{DAC}(s) \;=\; \frac{1-e^{-sT_s}}{s}V_o\,, \quad \text{NRZ DAC} \tag{1.13}$$

$$\;=\; \frac{1-e^{-s0.5T_s}}{s}2V_o\,, \quad \text{RZ DAC} \tag{1.14}$$

$$\;\approx\; \frac{\tau}{1+s\tau}V_{o,exp}\,, \quad \text{Exp. DAC} \tag{1.15}$$

where V_o is the output voltage of an NRZ DAC, $V_{o,exp}$ is the peak output of the exponential DAC and τ denotes the exponential decay constant with $\tau \ll 0.5T_s$ and $V_{o,exp}\tau = V_oT_s$.

The desired discrete-to-continuous time transformation can be done by solving eq. 1.12 with a symbolic analysis program like MAPLE. Useful transformation tables for modulators with NRZ DAC can be seen in Cherry and Snelgrove (1999*b*) and Shoaei (1996).

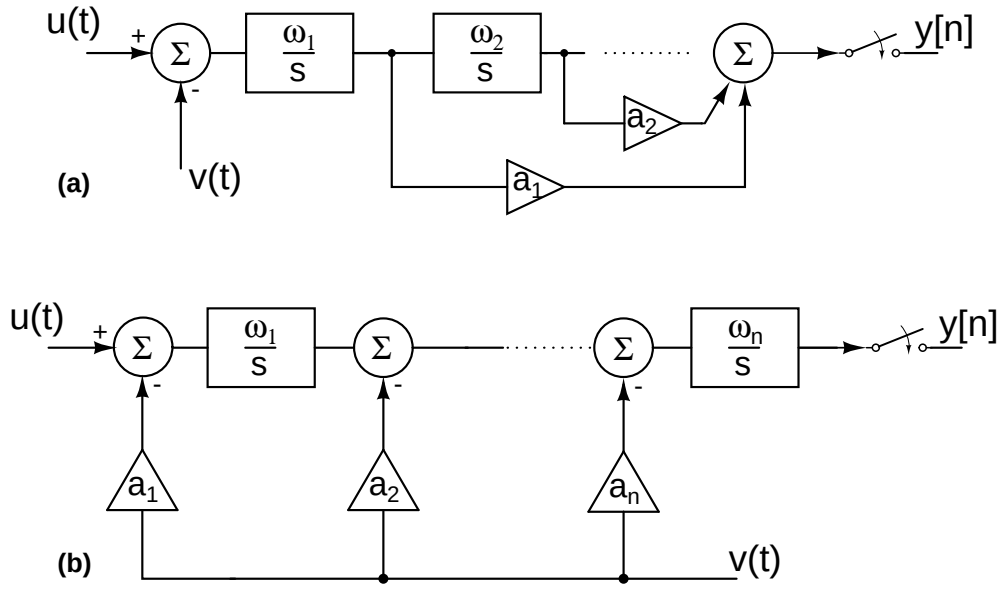


Figure 1.19: Implementation of the continuous time loop filter using (a) Cascade of Integrators Feedforward (CIFF) architecture and (b) Cascade of Integrators Feedback (CIFB) architecture

Alternatively, a numerical technique provided in Pavan (2010*b*) can be adopted to determine directly the loop filter coefficients that can realize $L_1(s)$ (and $L_0(s)$). Shown in Fig. 1.19 are the CIFF and CIFB loop filter architectures that can be used to implement the continuous-time loop filter. To illustrate the numerical method for the case of the CIFF loop filter, all the integrators can be assumed to have a gain of unity[13] ($\omega_i = 1$). The sampled responses of each integrator's path ($l_i[n]$) are then obtained through ideal block level simulation. From the knowledge of $l_i[n]$, the sampled response of the loop filter ($l[n]$) can be determined as

$$l[n] = \sum_{i=1}^{n} a_i l_i[n] \tag{1.16}$$

Defining $h[n]$ to be the impulse response obtained from the desired NTF, the solution

[13]For simplicity, the sampling rate, f_s is assumed to be 1 Hz

to the time domain equation[14],

$$h[n] + h[n] * l[n] = \delta[n] \tag{1.17}$$

helps in determining the coefficients $a_1, a_2, \cdots a_n$ and thereby realizing the loop filter. Since no assumptions need to be made on the DAC pulse shape, this technique can be used to determine the required loop filter, for any arbitrary DAC pulse shape.

The realized loop filter is then appropriately *node scaled* to ensure the swings at each of the integrator outputs are well within the limits set by the supply and the opamp swing constraints. *Frequency scaling* is then done to modify the loop filter suitable to be operated at the required sampling frequency. Note that such a frequency scaling is required as the impulse-invariant transformation (eq. 1.12) is specific to a given sampling rate (or sampling period).

The rest of the design process involves building the circuits necessary to implement the integrators, ADC and DAC. While the ADC is typically implemented with a flash architecture, the DAC can either have a resistive or current-steering topology. With regards to the integrators, several ways of implementation can be found in the literature, with each topology having its own merits and demerits. Fig. 1.20(a) shows the implementation of the first integrator of a CTDSM using the G_m-C architecture[15]. It can be seen that the transconductor's current gets integrated by the capacitor (C), thus realizing the transfer function $-\omega_1/s$, where $\omega_1 = g_m/2C$. Another popular technique widely used is the active-RC implementation, as shown in Fig. 1.20(b) with a resistive

[14]This is the time-domain equivalent of eq. 1.8, where '$*$' denotes the convolution operation.

[15]A resistive network has been used in this depiction to accomplish the summation of the input and fedback DAC signals. Alternately, a current-steering DAC can also directly drive the capacitor.

DAC. The high gain provided by the active circuit maintains virtual ground at v_d, thus enabling the current, $(u - v)/R$ to be integrated by the capacitor, C. Possible choices for the active-circuit include

- Single-stage transconductor
- Two-stage Miller compensated opamp (see Fig. 1.20(c))
- Two-stage feedforward compensated opamp (as shown in Fig. 1.20(d))

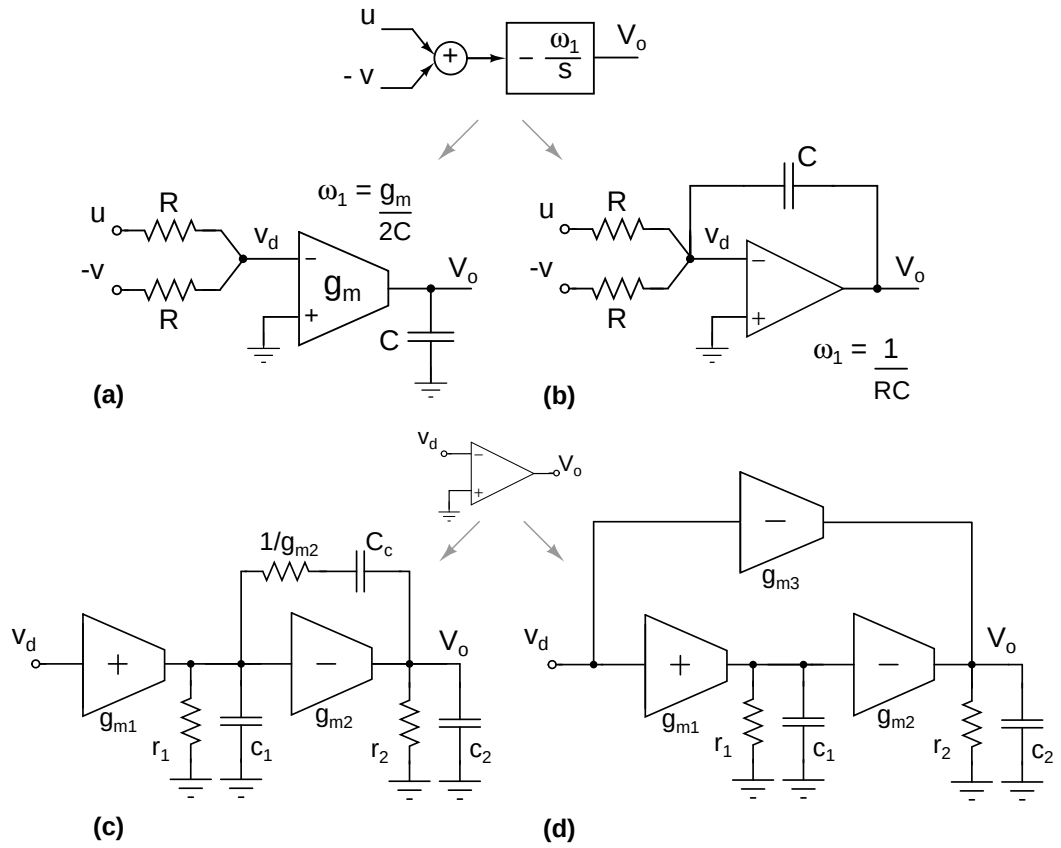


Figure 1.20: Input stage of a CTDSM implemented using (a) G_m-C integrator (b) active-RC integrator. Small-signal circuit models of (c) Miller compensated opamp (d) Feedforward compensated opamp

After the successful completion of the design process, the last and significant step involved in a CTDSM design is a process called *design centering*. The aim of this step is to tune the coefficients of the loop filter $(a_1, \cdots a_n)$, so that the desired NTF is obtained even in the presence of circuit nonidealities. The numerical technique as mentioned

before can be best used for this purpose. The reasons for deviation of the NTF include nonidealities like finite gain, finite bandwidth and parasitic poles of the opamps, which are used to implement the integrators. Another fundamental reason could be *excess loop delay*. It is defined as the delay between the sampling instant of the ADC and the time when the DAC is able to produce the pulse corresponding to the sample. It should be mentioned here that the problem of excess loop delay is one of the design issues specifically encountered with a CTDSM unlike its discrete-time counterpart.

1.4.3 Design issues in continuous-time modulators

Several issues arise when implementing continuous-time delta-sigma modulators. This section is aimed at providing a comprehensive overview of the works that have focused on these issues and in summarizing the various solutions that have been proposed.

A. Excess loop delay : Excess loop delay in CTDSMs has been investigated well by several authors over the years, as it can be a serious issue especially in high-speed CTDSMs. Due to excess loop delay (denoted by τ and modeled as $e^{-s\tau}$) the equivalence between the continuous-time and the discrete-time loop filter gets modified, as depicted in Fig. 1.21.

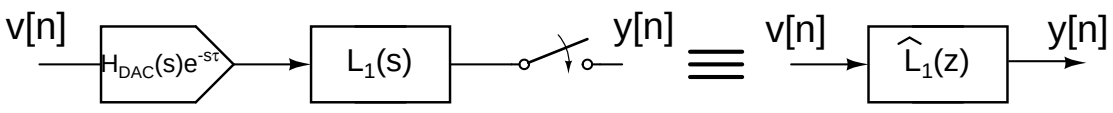


Figure 1.21: Equivalence between continuous-time and (the modified) discrete-time loop filter ($\widehat{L}_1(z)$) in the presence of delay (τ)

Cherry and Snelgrove (1999*b*) show that the modified loop filter ($\widehat{L}_1(z)$) can have an increased order with an NRZ DAC. As a consequence, the stability of the modulator is

reduced, thereby resulting in lesser MSAs with increasing loop delay. Shoaei (1996) illustrates this trend with the movement of the poles of NTF towards the unit circle, as the delay is increased. It is observed that modulators with lower OBG are more immune to the problem of excess loop delay.

The problem of delays upto half a clock cycle can be addressed by using RZ DACs. When using NRZ DACs, an intuitive solution is to bypass the slower path involving the feedback DAC and loop filter. This can be done by having a direct path around the ADC, through an additional DAC (Benabes *et al.* (1997)). A power efficient way of implementing this direct path can be found in Pavan *et al.* (2008). The desired NTF can as well be obtained by design centering the loop filter (Gao *et al.* (1997)).

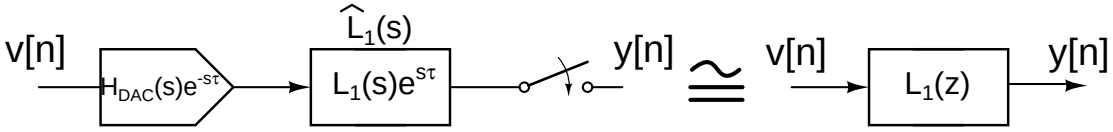


Figure 1.22: Conceptual model to compensate for excess loop delay

As conceptualized in Fig. 1.22, an intuitive technique for design centering has been provided by Pavan (2008), wherein the design centered loop filter is modeled as $L_1(s)e^{s\tau}$. Such a model results in a simpler method to determine the modified loop filter coefficients ($\widehat{a_1}, \widehat{a_2}, \cdots \widehat{a_n}$). A comparative study of all the existing compensation methods can be found in Keller *et al.* (2008). In short, the effect of excess loop delay has been well investigated and proven methods are available for addressing it.

B. Time constant variations : As mentioned earlier, the disadvantage with continuous-time implementation of the loop filter is the variation of time constants of the integrators (RC or C/g_m). Note that any variation in the loop filter transfer function, $L_1(s)$ modifies the equivalent discrete-time loop filter and correspondingly results in a modified

NTF. To elaborate, let k_p denote the relative change in the time constants of the integrators, that results in a modified loop filter transfer function[16], $L_1(s/k_p)$. Assuming a third order NTF with an OBG of 3 (and OSR of 64), the observed variation in the NTF is illustrated in Fig. 1.23, for extreme cases of k_p (1.3 and 0.7).

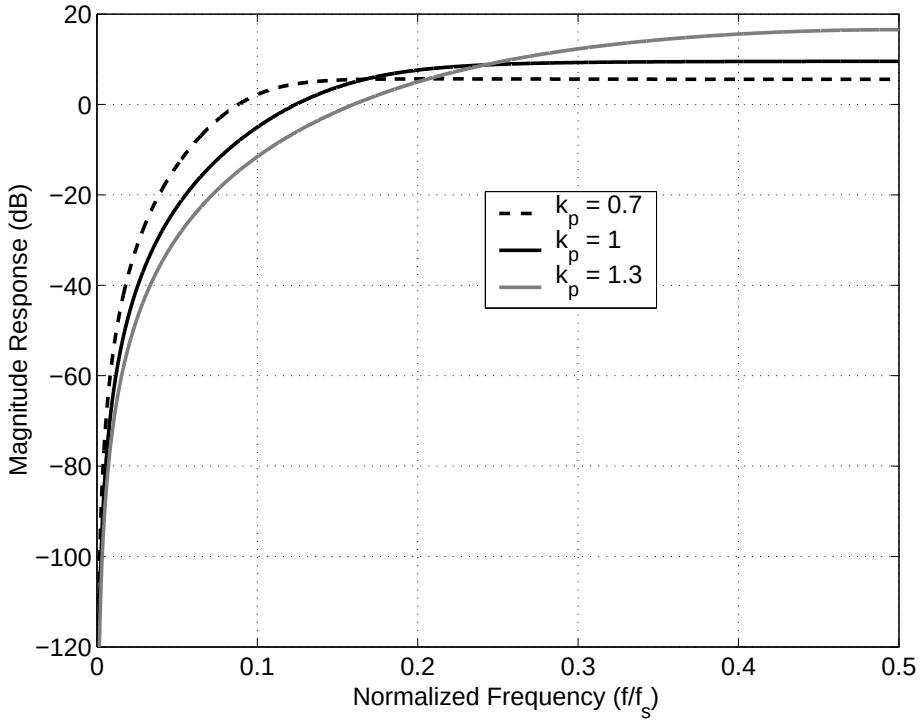


Figure 1.23: Magnitude response of the NTF for different values of k_p

When $k_p > 1$, the unity gain frequency of the loop filter is increased, which in-turn pushes the poles in the NTF to higher frequencies. Such a movement of the poles can be expected to result in a higher out-of-band gain, with a corresponding improvement in the in-band noise shaping. While on one hand this can be beneficial due to lesser in-band quantization noise, on the other, it can drastically reduce the MSA of the modulator due to an increased OBG. Equivalently, the case of $k_p < 1$ leads to an increased stability of the modulator at the expense of an increased in-band noise. Thus, it can be seen that both extremes of variations in time constants can cause problems in the performance of CTDSMs.

[16]assuming the variation is the same in all the integrators implementing the loop filter

Parallels can be drawn at this juncture between the time constant variations and the variability in the sampling rate. It can be seen that reducing the sampling rate is equivalent to a reduction in the time constants ($k_p > 1$) of the loop filter, akin to frequency scaling. Thus, CTDSMs with a given loop filter become unstable when the sampling rate is reduced beyond a limit. More importantly, it is for this reason that CTDSMs have satisfactory operation only over a narrow range of sampling rates.

The conventional way of addressing the problem of time constant variations is to tune the time constants to make sure they are within tolerable limits. The entire process involves the determination of the variation in the time constants, followed by relevant tuning of the loop filter to bring back the time constants closer to the nominal values. Several methods have been proposed to automate this procedure. The authors in Xia *et al.* (2004) make use of a replica oscillator to estimate the time constant on-chip and apply the required tuning to the modulator's loop filter. Loke *et al.* (2005) utilizes the knowledge of the gain of a replica integrator at a given frequency to fix the tuning required in the modulator. An 'in-situ' or direct tuning method has been devised by Pavan and Krishnapura (2007), where they make use of a digital technique to determine the variance of $v[n] - v[n-1]$ and use it as a measure to estimate the variation in time constants. Since the technique is largely developed on a linearized model for the quantizer, it is found to have limited accuracy when quantizer levels fewer than four are used. A tuning technique that can be applied to single-bit modulators has been proposed by Saxena *et al.* (2009). It is based on measuring the variance of the output of the first integrator, to determine the change in the time constants. In summary, the problem of time constant variations can be taken care by choosing a suitable tuning algorithm, to get back the time constants to within tolerable limits.

C. DAC clock jitter : One of the advantages of CTDSMs is that they can tolerate the timing errors that occur at the input of the ADC, as they get shaped by the NTF. However, the same tolerance cannot be found with timing errors in the DAC, as they appear directly at the input of the modulator. These timing errors can be either pulse-width errors (random or accumulated jitter) or pulse-delay errors. Pulse-delay jitter can be an issue with RZ DACs. However, it does not affect the modulator performance as much as pulse-width jitter, since the error is first-order noise shaped (Oliaei (1999)).

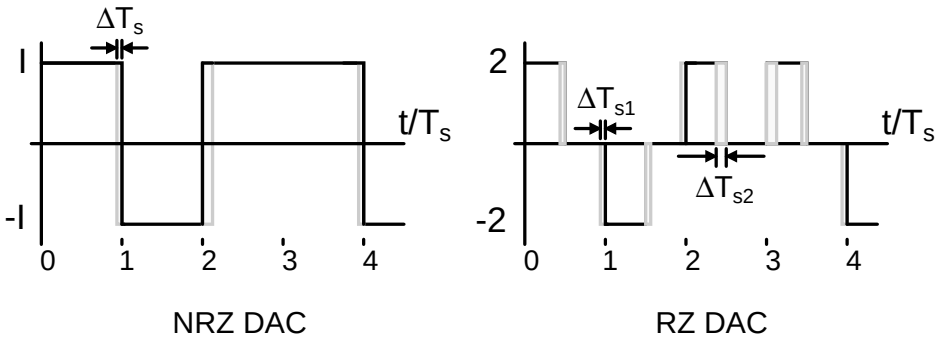


Figure 1.24: Pulse width timing errors seen in NRZ and RZ DAC waveforms of a single-bit CTDSM

The effect of pulse-width jitter (as shown in Fig. 1.24) is to change the effective pulse-widths of the feedback DAC waveforms in each clock cycle. Since such errors get integrated by the following loop filter, the effect on the modulator performance is determined by 'averaging' the error introduced in each clock cycle. The averaged error sequence thus obtained, can be modeled as an additional input to a jitter-free modulator and is given by

$$e_j[n] = (v[n] - v[n-1])\frac{\Delta T_s}{T_s}, \qquad \text{NRZ DAC} \tag{1.18}$$

$$= 2v[n]\left(\frac{\Delta T_{s1}}{T_s} + \frac{\Delta T_{s2}}{T_s}\right), \qquad \text{RZ DAC} \tag{1.19}$$

where ΔT_{s1} and ΔT_{s2} are the respective timing errors of the two edges in a clock pe-

riod, as shown in Fig. 1.24. Interpreting the above equations in the frequency domain, the spectrum of the additional random input ($e_j[n]$) can be seen as resulting from the 'modulation' of the output sequence by the jitter noise sequence. Due to the presence of high out-of-band components in the output spectrum, this can lead to increased in-band noise for the CTDSM. Cherry and Snelgrove (1999*a*) were among the first to study this effect of clock jitter on the performance of single-bit CTDSMs. They concluded that modulators with NRZ DAC can have better performance than with an RZ DAC. As evident from eq. 1.18 and 1.19, this can be attributed to the dependency of the jitter noise sequence on the timing errors of both the clock edges (ΔT_{s1} and ΔT_{s2}), in case of the RZ DAC. It is also due to the fact that an error is introduced with an NRZ DAC, only when there are transitions ($v[n] - v[n-1]$) in the modulator output. This is in contrast to an RZ DAC where transitions occur every clock cycle.

Observing that the demodulation of the out-of-band noise is a key contributor to the performance degradation due to clock jitter, Hernandez *et al.* (2004) pointed out that the effect of clock jitter can be minimized by optimizing the shape of the NTF around $f_s/2$. This was later analyzed mathematically by Reddy and Pavan (2007) through the use of Bode's sensitivity integral. It was shown that a trade-off is required while choosing OBG, as modulators with higher OBG are more prone to the effects of clock jitter.

Several ways have been proposed to counter the issue of clock jitter. An obvious solution as evident from eq. 1.18 is to increase the number of levels in the quantizer. This reduces the height of the transitions occurring in each clock cycle, thereby resulting in better immunity to clock jitter. Since the use of multi-bit quantizers necessitates the use of Dynamic Element Matching (DEM) techniques (to shape away the DAC

mismatch errors)[17], other solutions have been proposed to overcome the effect of clock jitter in single-bit modulators.

One such solution is to use an exponentially decaying DAC pulse (Ortmanns *et al.* (2005)), generated as shown in Fig. 1.25. This type of DAC pulse generator has been referred to as the Switched-Capacitor-Resistor[18] (SCR) DAC in the literature.

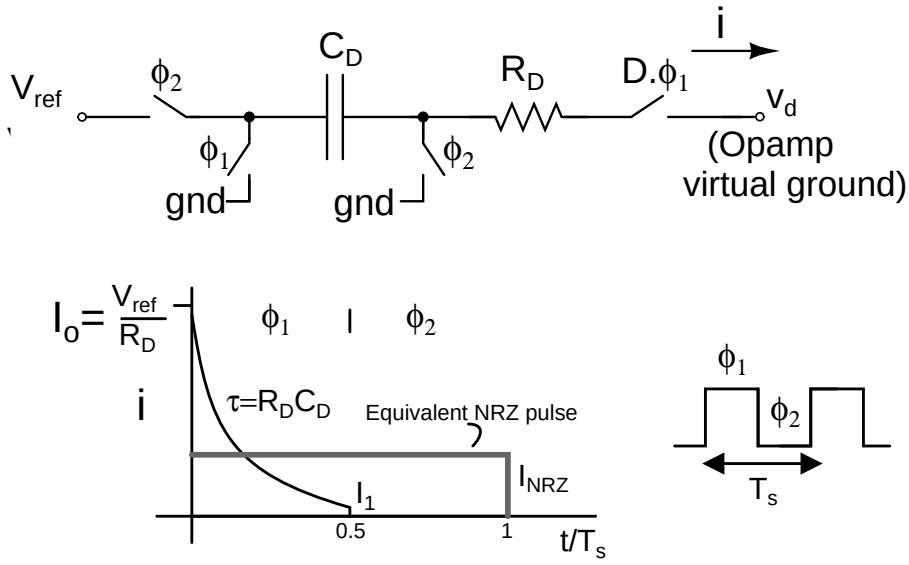


Figure 1.25: SCR DAC schematic and the exponentially decaying pulse produced by the DAC

To briefly describe its operation - in clock phase ϕ_1, the capacitor C_D (which has been previously charged to V_{ref} during ϕ_2) discharges into the virtual ground of the opamp through a resistor R_D with a time-constant $\tau = C_D R_D$. The initial (peak) current injected by the DAC is $I_o = V_{ref}/R_D$, decaying to $I_1 = I_0 \exp(-T_s/(2\tau))$. If $I_1 << I_0$, the average current is given by $C_D V_{ref} f_s$ (the height of the equivalent NRZ DAC pulse) and $I_1 < C_D V_{ref} f_s$. With this DAC, the noise introduced due to clock jitter is proportional to I_1 and is controlled by T_s/τ. Smaller the value of τ (small R_D or C_D), smaller is the value of I_1 and thus, smaller is the sensitivity to clock jitter.

[17]Blocks such as DEM in the feedback path can increase the excess loop delay of the modulator, thus creating issues in high-speed CTDSMs. The use of the inherently linear single-bit quantizer can alleviate such delay problems.

[18]The given implementation assumes that the first integrator is implemented with active-RC technique

An alternative solution to the SCR DAC is to use a FIR DAC in the feedback path (as in Oliaei (2003) and Putter (2004)), to convert the output of the single-bit ADC to multi-level signals. Thanks to the multi-level nature thus obtained, the system is more tolerant to the effects of clock jitter. Very recently, Colodro and Torralba (2009) have applied this concept to a multi-bit modulator. They convert the internal multi-bit ADC output to a single-bit pulse-width modulated signal and feed it back to the loop filter through an FIR DAC. In doing so, they show that the same immunity to clock jitter as provided by a multi-bit DAC can be achieved, without the need for DEM techniques. In summary, similar to the excess loop delay, several techniques have been proposed for mitigating the effects of DAC clock jitter.

1.4.4 Issue of integrator nonlinearity in continuous-time modulators

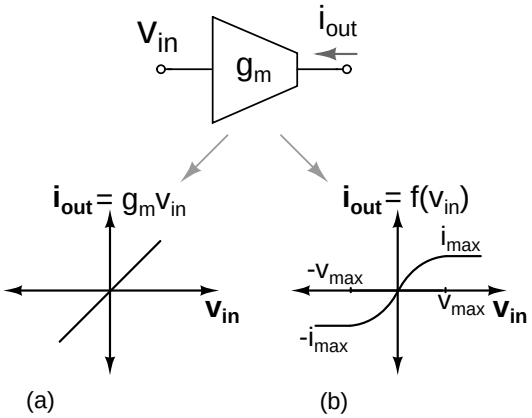


Figure 1.26: Input-output characteristic of (a) a linear transconductor and (b) a nonlinear transconductor

Another nonideality that is critical to the performance of CTDSMs, is the nonlinearity of integrators. An important source of this nonlinearity are the *transconductors* (g_m) used to implement the opamps in the integrators. While an ideal transconductor exhibits a linear relation between its output current (i_{out}) and its input voltage (as in Fig. 1.26(a)), a practical transconductor is observed to have a nonlinear relation, $f(v_{in})$.

With a differential implementation, the nonlinear relation[19] can be expressed as

$$i_{out} \;=\; f(v_{in}) \;=\; g_m v_{in} - g_3 v_{in}^3\,, \qquad (|v_{in}| \,<\, v_{max}) \tag{1.20}$$

Further, for very high input voltages ($> v_{max}$), the output current of the transconductor invariably gets limited to a value of i_{max}, as can be seen in Fig. 1.26(b)[20]. When the output current is saturated, the transconductor is said to be operating in a 'slewing' mode. In a CTDSM, slewing can result in errors in the integrator output, leading to performance degradations in the form of increase in the in-band noise and harmonic distortion in the modulator output (Cherry (1998)).

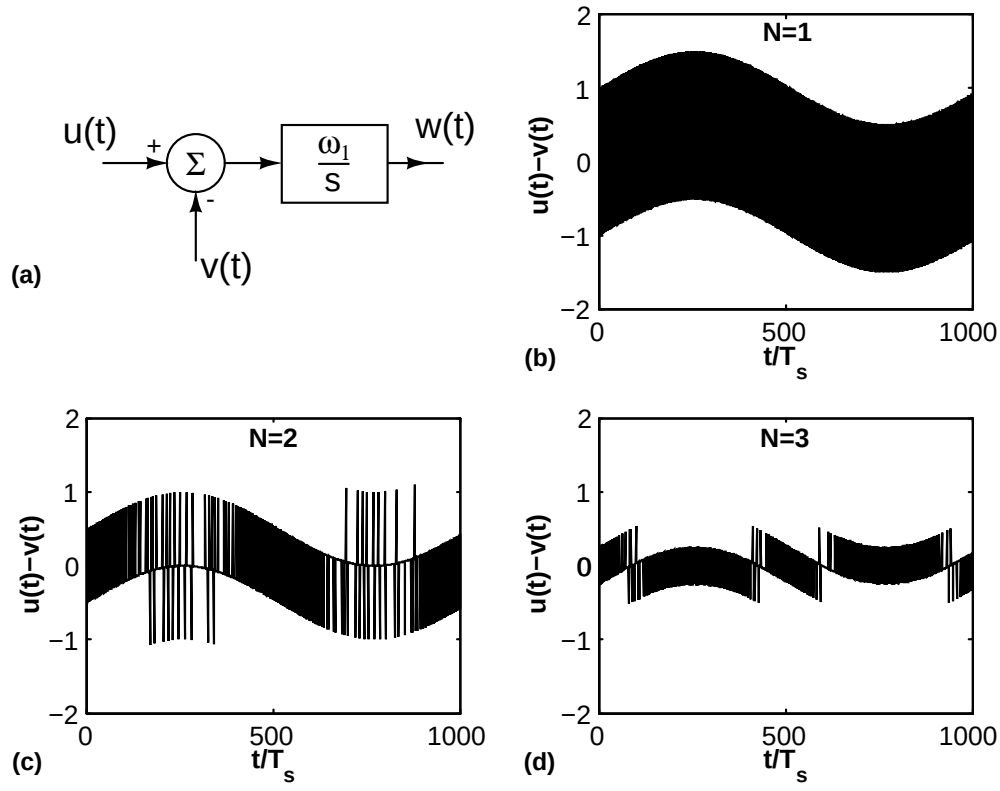


Figure 1.27: (a) Block diagram of the input stage of a CTDSM including the first integrator, along with typical swings observed for u(t)-v(t) for different number of quantizer bits (b) N = 1 (c) N = 2 (d) N = 3. (Full scale range of the quantizer is 2)

[19]We restrict our attention to third order nonlinearities

[20]For instance, in a typical differential pair implementation of the transconductor (Razavi (2005)), i_{max} is given by the tail current used.

Therefore, the primary task while designing integrators for CTDSM is to ensure that the specification called Slew-Rate (SR) requirement is satisfied. SR is defined as the maximum rate of change of the output voltage possible in a given circuit, with a given transconductor. In case of the first integrator of a CTDSM (as modeled in Fig. 1.27(a)), the slew-rate needs to be greater than the expected rate change of voltage at the first integrator output ($w(t)$). Thus, we have

$$SR \geq \left|\frac{d\,w(t)}{dt}\right|_{max} = \left|\frac{d}{dt}\left(\omega_1 \int_0^t (u(t_1) - v(t_1))\; dt_1\right)\right|_{max} \tag{1.21}$$

$$= \omega_1 \left|u(t) - v(t)\right|_{max} \tag{1.22}$$

The above equation indicates that the slew-rate requirement is controlled by the 'peak' of $\left|u(t) - v(t)\right|$. Recall that SCR DAC implementations have very high initial peak outputs, mainly intended to enhance the immunity to clock jitter. This translates to higher slew-rate requirements imposed on the first integrator of the CTDSM. For the same reason, such requirements are also higher with an RZ DAC[21], when compared to its NRZ counterpart. While a brute-force way of avoiding slewing conditions is to increase the current sourcing capabilities of the transconductors, it can be seen that an alternate power efficient design choice is to adopt NRZ pulse shapes.

Given an NRZ implementation, power consumption of the first integrator can be further minimized through proper choice of the number of quantizer levels. As can be seen from Fig. 1.27(b), (c) and (d), increasing the number of quantizer bits can help in reducing the peak value of $\left|u(t) - v(t)\right|$ and thereby the slew-rate requirement. Further, as seen from eq. 1.22, an appropriately smaller ω_1 (achieved through node scaling) can also help in controlling the slew-rate requirements placed on the first integrator. In a

[21] since the height of v(t) is twice of that with an NRZ DAC

nutshell, the critical issue of slewing can be avoided and power consumption can be optimized through proper architectural design choices.

Even when the current levels are high enough to prevent slewing in the integrators, the performance of a CTDSM is not guaranteed. This is due to the nonlinear nature of the transconductor's current, as defined by $f(v_{in})$. Of particular interest is the effect of the transconductor operating in a weakly nonlinear region, where $g_m v_{in} \gg g_3 v_{in}^3$. Referred to as opamp nonlinearity or integrator nonlinearity henceforth, such a nonideality increases the current requirements of the opamp[22]. This makes integrator nonlinearity a critical consideration while designing low power, high resolution CTDSMs. To demonstrate this, consider a CTDSM with a 4-bit internal quantizer and a third order maximally flat NTF. A sampling rate of 3.072 MHz is chosen which translates to an OSR of 64 for a signal bandwidth of 24 kHz. As shown in Fig. 1.28, let the first integrator be implemented with a nonlinear G_m-C integrator. The transconductor is assumed to have a current function, $i_{out} = g_m v_d - g_3 v_d^3$. The nonlinearity is such that the transconductor is ensured to operate in the weakly nonlinear region ($g_m v_d \gg g_3 v_d^3$).

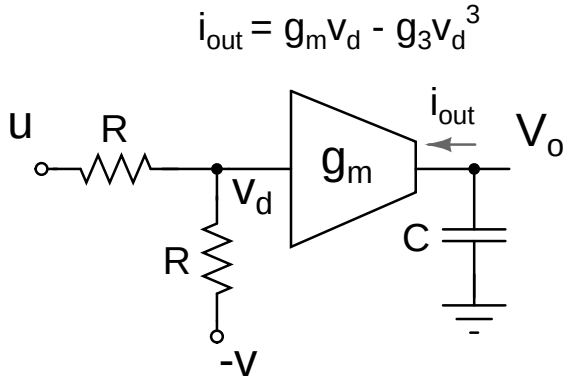


Figure 1.28: A nonlinear G_m-C implementation of the first integrator

Despite the fact that the transconductor never slewed, a significant increase in the in-

[22] even more than that is sufficient to avoid slewing

band noise floor (when compared to a linear modulator) can be seen from the Power Spectral Density (PSD) of the modulator output shown in Fig. 1.29.

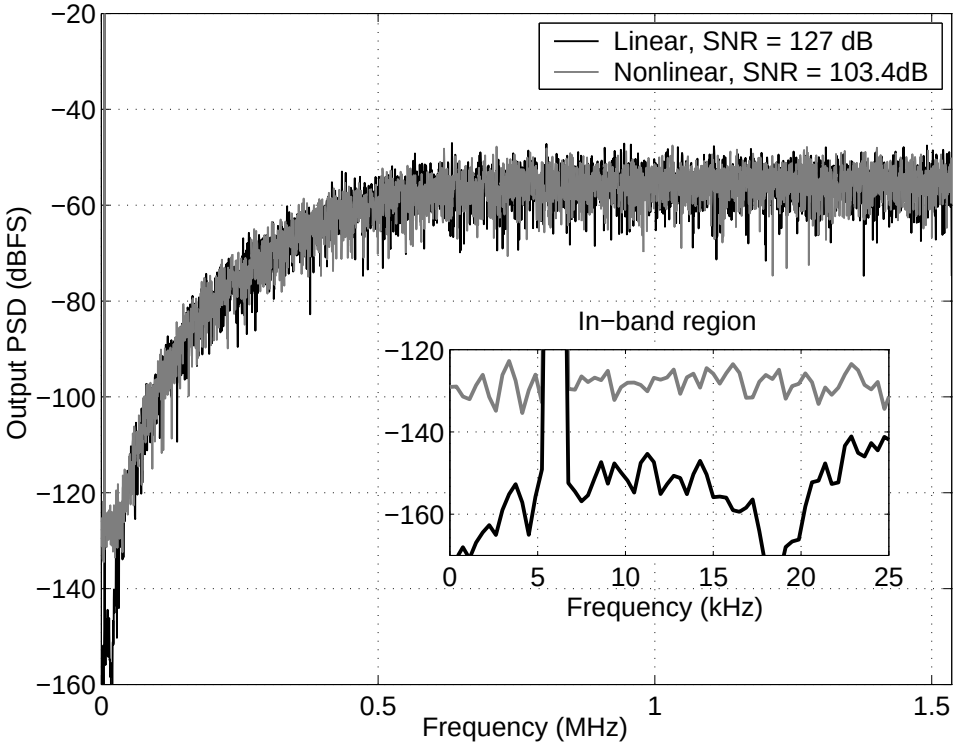


Figure 1.29: Simulated PSD of a third order CTDSM with and without nonlinearity in the G_m-C integrator. Comparison of the in-band noise floor is shown in the inset.

An interesting observation is that, for the same nonlinearity, the in-band noise increases with increasing OBG, as shown in Fig. 1.30. Similarly, it is observed that a 3-bit modulator is affected more by integrator nonlinearity than a 4-bit modulator (as shown in Fig. 1.31). Recognizing the trend of increasing in-band noise floor with reducing number of quantizer levels, the modulator was simulated with a single-bit quantizer[23] to observe the performance degradation. Fig. 1.32 shows the PSDs of the modulator output observed for two-different input amplitudes. Unlike its multi-bit counterpart, the performance of the single-bit modulator is seen to be largely dented by harmonic distortion observed in its output.

[23]For this simulation, the modulator was modified to have a maximally flat NTF with an OBG of 1.5. Also, an OSR of 128 was used resulting in a sampling rate of 6.144 MHz.

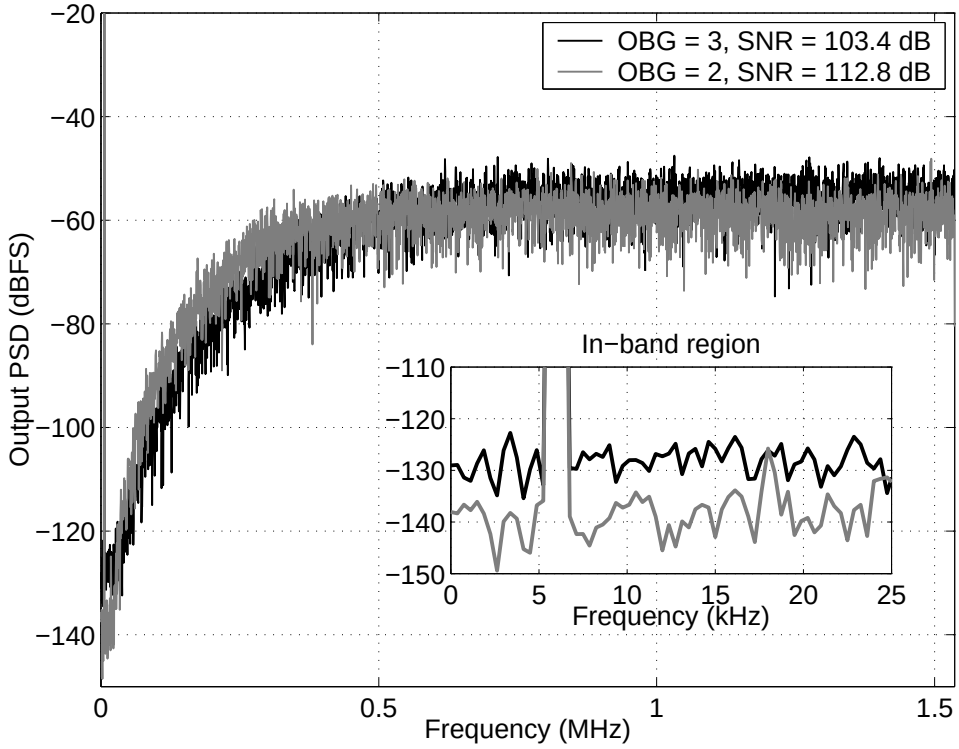


Figure 1.30: Simulated PSD of a third order CTDSM with a nonlinear transconductor for different OBGs.

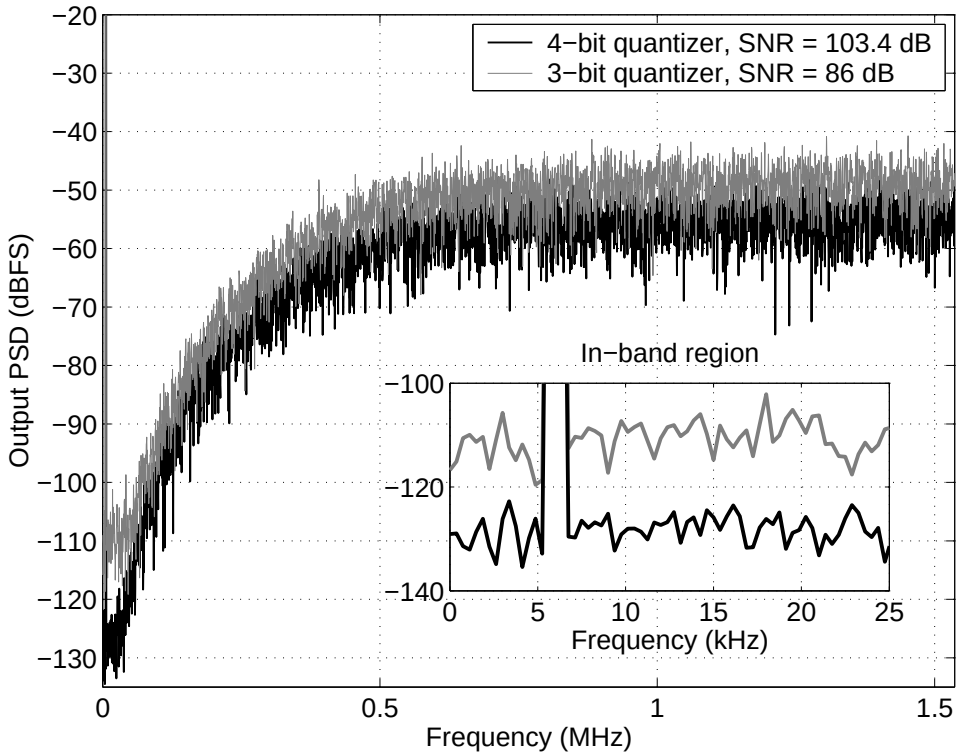


Figure 1.31: Simulated PSD of a third order CTDSM with a nonlinear transconductor for different quantizer levels

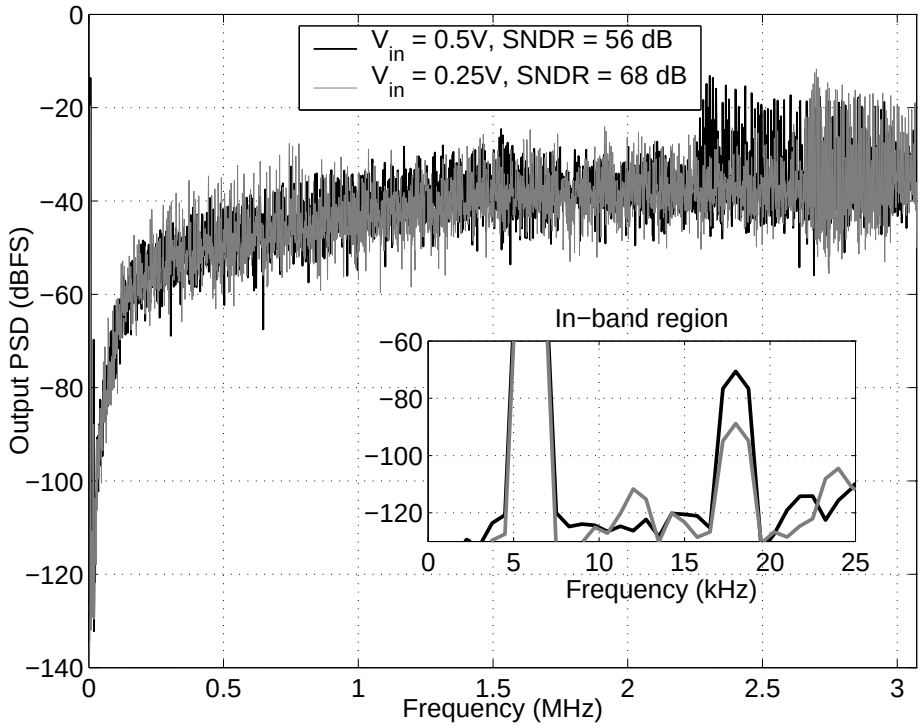


Figure 1.32: Simulated PSD of a third order single-bit CTDSM showing the harmonic distortion limited performance for two different input amplitudes

In summary, the severity of integrator nonlinearity on the performance of CTDSMs should be evident from the above discussion.

1.5 Prior work and Motivation

A survey on the studies that have been carried out on analyzing the effect of integrator nonlinearity in CTDSMs, points to the works of Breems *et al.* (1999) and Leuciuc (2001). Breems *et al.* (1999) quantified the harmonic distortion observed in single-bit CTDSMs (with an RZ DAC) employing a single-stage active-RC integrator. Leuciuc (2001) has shown through behavioral simulations the relative significance of integrator nonlinearities. He concludes that the first integrator of the loop filter is the dominant contributor to harmonic distortion, in both CIFF and CIFB modulators.

On closer observation, it is seen that the works mentioned above have focused on

particular integrator topologies and DAC pulse shapes, leaving much to be investigated for other design choices. Moreover, to the author's best knowledge, there has been no work explaining the increase in the in-band noise of multi-bit modulators, due to integrator nonlinearity[24].

Due to the limited understanding of the effect of integrator nonlinearity, designers have to depend on time-consuming simulations to determine the power required to meet the performance specifications. This often involves a "trial and error" approach of increasing the bias currents of the transconductors, till the desired performance is achieved. Further, the power efficiency achieved with such designs is debatable, given the lack of proper understanding of this issue.

From our discussions so far, it must be clear that all other implementation issues of a CTDSM have been well addressed except for the effect of integrator nonlinearity. From a designer's perspective, a model that can quantify the performance degradation due to such nonlinearities can be very useful in designing low power, high-resolution CTDSMs. A thorough understanding of the issue can also throw light on methods that can mitigate the observed performance degradations. This research work is aimed at addressing the above mentioned needs.

1.6 Contributions

Two major contributions of this research are the following. The effect of nonlinearity in the first integrator of a CTDSM has been analyzed, when a multi-bit quantizer is used. Analytical expressions have been developed to quantify the increase in the

[24] A similar phenomenon was briefly pointed out by Steensgaard (1998) for 'single-bit discrete-time' modulators

in-band noise (IBN) spectral density, relating it the nonlinearity in several integrator topologies. Such relations have been obtained by assuming the shaped quantization noise to be approximately Gaussian. Results from macromodel simulations have been found to correlate well with the theory.

Another contribution is a power efficient technique to achieve low distortion operation in CTDSMs, especially with single-bit quantizers. Termed as the "assisted opamp technique", its efficacy has been demonstrated through two single-bit audio modulator designs - one using an NRZ feedback DAC and the other using an SCR feedback DAC. With the help of the proposed technique, the performance achieved with these modulators is comparable to that obtained with the best multi-bit designs reported in the literature. Designed in a 0.18 μm CMOS technology, the modulator with an NRZ DAC achieves a dynamic range of 92.5 dB in a 24 kHz bandwidth, with a power dissipation of 110 μW from a 1.8 V supply. The design employing an SCR DAC achieves a dynamic range of 91.5 dB and dissipates 122 μW of power. A detailed description of these contributions along with details of the fabricated designs form the rest of the thesis, which has been organized as described below.

1.7 Organization of the thesis

Chapter 2 focuses on the analysis of the effect of integrator nonlinearity in multi-bit CTDSMs. Analytical relations quantifying the increase in the IBN of multi-bit modulators will be derived for different integrator topologies, considering both NRZ and RZ feedback DACs.

Chapter 3 builds on the fundamentals developed in the previous chapter to analyze

the effect of integrator nonlinearity in specific active-RC integrator topologies. It shows how the models derived in Chapter 2 can be suitably modified to be applied to these integrator topologies.

Chapter 4 discusses as to how the concepts derived in the Chapters 2 and 3 can be applied to determine the harmonic distortion observed in single-bit modulators.

Chapter 5 introduces the circuit technique of assisted opamp integrator that can enable power efficient design of CTDSMs, especially single-bit modulators.

Chapter 6 then provides the design details of the two audio modulators designed to demonstrate the effectiveness of the assisted opamp technique. The relevant simulation and measurement results will also be provided in this chapter.

Chapter 7 summarizes the contributions and the results achieved through this research and also gives suggestions for future work.

CHAPTER 2

ANALYSIS OF INTEGRATOR NONLINEARITY IN CONTINUOUS-TIME DSMs : THE MULTI-BIT CASE

The effect of integrator nonlinearity has been shown to increase the in-band noise floor of multi-bit continuous-time delta-sigma modulators. In this chapter[1], we analyze this effect and show how it can be related to the NTF of the modulator and the number of quantizer levels, among other parameters. The analysis will be carried out for a CTDSM having a CIFF loop filter as shown in Fig. 2.1(a)[2]. Since the first integrator is the dominant source of nonlinearity (Leuciuc (2001)), the rest of the loop filter $L'(s)$ is considered to be linear. Without loss of generality, the DAC is assumed to be of the resistive kind, thus resulting in the input stage as shown in Fig. 2.1(b).

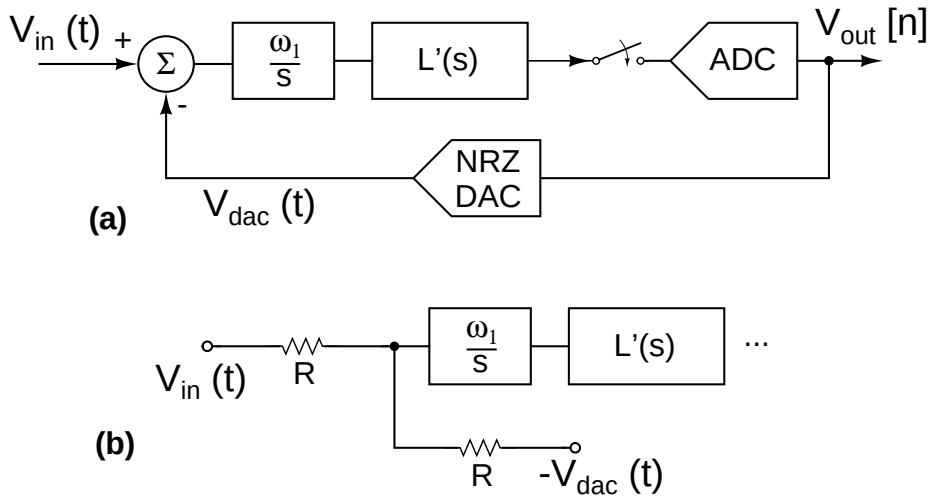


Figure 2.1: (a) Block diagram of a modulator with CIFF loop filter and (b) the input stage with a resistive DAC implementation.

[1]The contents of this chapter have appeared (in condensed form) in Sankar and Pavan (2007).

[2]In the rest of the thesis, the input signal will be denoted as $V_{in}(t)$ instead of $u(t)$.

Considering an NRZ feedback DAC, we proceed with the analysis in the following manner. First it is shown how a unified model can be used to account for the nonlinearity in integrators employing various opamp topologies. Using this model, analytical expressions are derived for the in-band noise of multi-bit modulators, relating it to the shape of the NTF. The analysis is then extended to modulators employing RZ feedback DACs. The efficacy of the proposed models is finally verified by comparing the analytical expressions with behavioral simulation results.

2.1 Modeling nonlinearity in G_m-C and active-RC integrators

In this section, the weak nonlinearity in integrators implemented using G_m-C and active-RC techniques will be modeled. The discussion has been structured in a way to bring about the best suited integrator topology that can be employed to design power efficient, linear integrators for use in CTDSMs.

2.1.1 G_m-C integrator

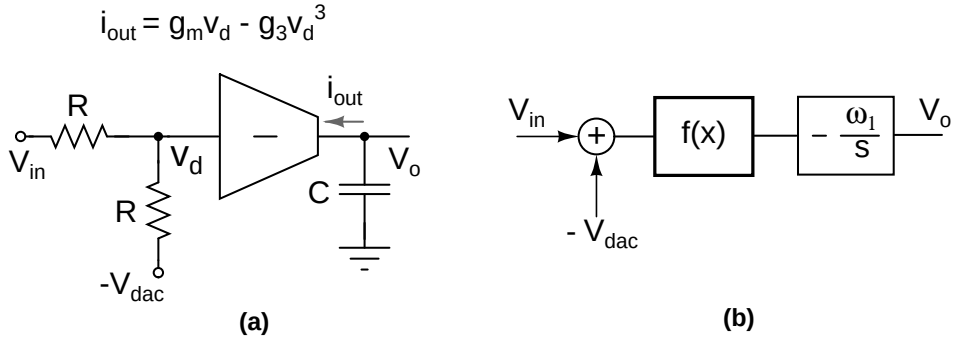


Figure 2.2: (a) Circuit of the input stage of a CTDSM using a nonlinear G_m-C integrator and (b) its equivalent model.

The input stage of a CTDSM using a G_m-C integrator is shown in Fig. 2.2(a). The transconductor is assumed to be weakly nonlinear, with the output current being related to the input voltage as $i_{out} = g_m v_d - g_3 v_d^3$. The output voltage of the nonlinear integrator

is

$$V_o = -\frac{1}{C}\int_0^t i_{out}\ \mathrm{dt} \tag{2.1}$$

$$= -\frac{1}{C}\int_0^t (g_m v_d - g_3 v_d^3)\ \mathrm{dt} \tag{2.2}$$

With $v_d = (V_{in} - V_{dac})/2$, we have

$$V_o = \frac{-1}{C}\int_0^t \frac{g_m}{2}(V_{in} - V_{dac}) - \frac{g_3}{8}(V_{in} - V_{dac})^3\ \mathrm{dt} \tag{2.3}$$

$$= \frac{-g_m}{2C}\int_0^t (V_{in} - V_{dac}) - \frac{g_3}{4g_m}(V_{in} - V_{dac})^3\ \mathrm{dt} \tag{2.4}$$

Thus, as shown in Fig. 2.2(b), the nonlinear integrator can be modeled with a nonlinear function preceding a linear integrator, where

$$f(x) = x - \frac{g_3}{4g_m}x^3 \tag{2.5}$$

$$\omega_1 = \frac{g_m}{2C} \tag{2.6}$$

2.1.2 Active-RC integrators

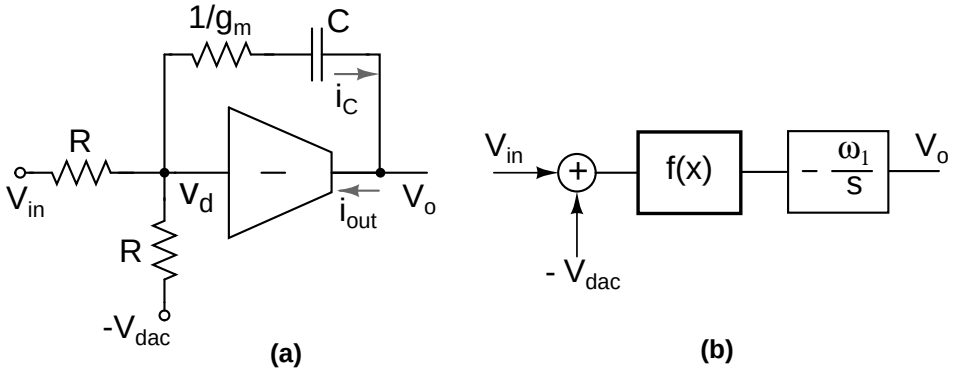


Figure 2.3: (a) Circuit of the input stage of a CTDSM using a nonlinear single-stage active-RC integrator and (b) its equivalent model.

1) Active-RC integrator using a transconductor: A popular CTDSM input stage using a zero-compensated active-RC integrator is shown in Fig. 2.3(a) (see for

example, van der Zwan and Dijkmans (1996), Breems *et al.* (1999), Yan and Sanchez-Sinencio (2004), Breems *et al.* (2004)). The transconductor is weakly nonlinear with $i_{out} = g_m v_d - g_3 v_d^3$. The output voltage of the integrator can be determined by solving the nonlinear equation involving v_d. The equation in question can be formed by writing the nodal equation at the transconductor input,

$$\frac{V_{in} - V_{dac}}{R} - \frac{2v_d}{R} = i_C = g_m v_d - g_3 v_d^3 \tag{2.7}$$

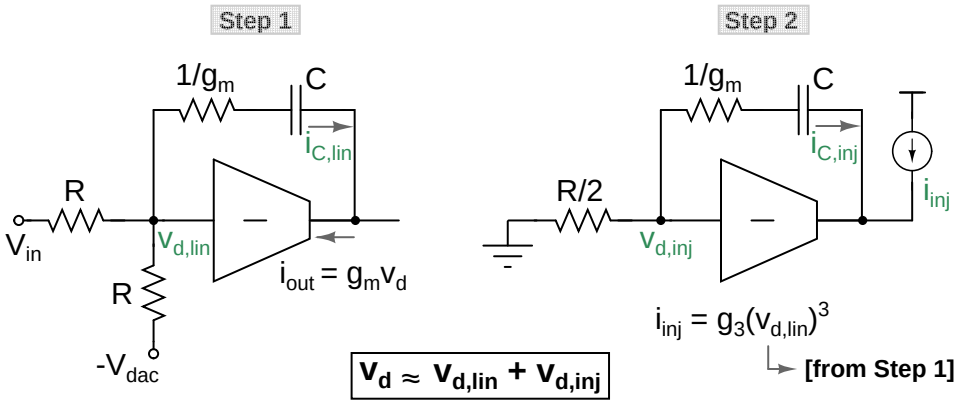


Figure 2.4: Circuit topologies in each step of the current injection method

The solution (v_d) for this equation can be obtained using the method of current injection[3] (for weakly nonlinear systems) proposed by Bussgang *et al.* (1974) as follows,

- **Step 1:** First, a solution ($v_{d,lin}$) is obtained by assuming a linear transconductor ($g_3 = 0$), as shown in Fig. 2.4. We thus have

$$v_{d,lin} = \frac{(V_{in} - V_{dac})}{2 + g_m R} \tag{2.8}$$

- **Step 2:** The next step is to find a solution for v_d with the same linear transconductor but after nullifying the input and 'injecting' a current $g_3 v_{d,lin}^3$ as shown in Fig. 2.4. Termed as $v_{d,inj}$, it can be given by

$$v_{d,inj} = \frac{R}{(2 + g_m R)} i_{inj} \tag{2.9}$$

[3]The reader can refer to Appendix A for a detailed description about the method. While the method is originally intended to solve nonlinear differential equations, it has been recently applied to efficiently simulate delta-sigma modulators (Pavan (2010*a*)).

- The solution to the nonlinear equation is approximately the sum of the two components thus derived.

$$v_d \approx v_{d,lin} + v_{d,inj} = v_{d,lin} + \frac{g_3 v_{d,lin}^3 R}{(2 + g_m R)} \tag{2.10}$$

The capacitor current (i_C) can now be determined using eq. 2.7 or equivalently,

$$\begin{aligned} i_C &\approx i_{C,lin} + i_{C,inj} = g_m v_{d,lin} - \frac{2 v_{d,inj}}{R} && (2.11)\\ &= g_m v_{d,lin} - \frac{2 g_3 v_{d,lin}^3}{(2 + g_m R)} && (2.12)\\ &= \frac{g_m (V_{in} - V_{dac})}{2 + g_m R} - \frac{2 g_3 (V_{in} - V_{dac})^3}{(2 + g_m R)^4} && (2.13) \end{aligned}$$

Expressing the output voltage as a function of i_C, we have

$$\begin{aligned} V_o &= -\frac{1}{C} \int_0^t i_C \, \mathrm{dt} \; - \frac{i_C}{g_m} + v_d && (2.14)\\ &= -\frac{1}{C} \int_0^t i_C \, \mathrm{dt} \; + \frac{g_3 v_d^3}{g_m} && (2.15)\\ &\approx -\frac{1}{C} \int_0^t i_C \, \mathrm{dt}\,, \qquad (\text{since } g_m v_d \gg g_3 v_d^3) && (2.16)\\ &= \frac{-1}{C} \int_0^t \left(\frac{g_m (V_{in} - V_{dac})}{2 + g_m R} - \frac{2 g_3 \left(V_{in} - V_{dac}\right)^3}{(2 + g_m R)^4} \right) \mathrm{dt} && (2.17) \end{aligned}$$

From the above equation, it is apparent that the integrator can be modeled by a nonlinear function ($f(x)$) preceding a linear integrator as shown in Fig. 2.3(b), with

$$\begin{aligned} f(x) &= x - \frac{2 g_3}{g_m (2 + g_m R)^3} x^3 && (2.18)\\ \omega_1 &= \frac{g_m}{C(2 + g_m R)} && (2.19) \end{aligned}$$

When compared to eq. 2.5, observe that the negative feedback around the transconductor has effectively reduced the nonlinearity by a factor $(2 + g_m R)^3$, where $g_m R/2$

denotes the loop gain. Physically, the improvement can be attributed to the reduced swing at the transconductor input (see eq. 2.8), which correspondingly reduces the injected nonlinear currents. The linearity of a given active-RC integrator can thus be seen to be improved by either increasing the transconductance (g_m) or by increasing the resistance, R. While the former choice leads to an increased power consumption, the latter increases the thermal noise contribution of the resistors.

2) Active-RC integrator with a two-stage Miller opamp: The input stage of a CTDSM with an active-RC integrator built with a two-stage Miller compensated opamp is shown in Fig. 2.5. The input transconductor is assumed to be weakly nonlinear, with $i_1 = g_{m1}v_d - g_{31}v_d^3$. The second stage is assumed to be a linear transconductor, g_{m2}. c_1 and c_2 denote the parasitic capacitances.

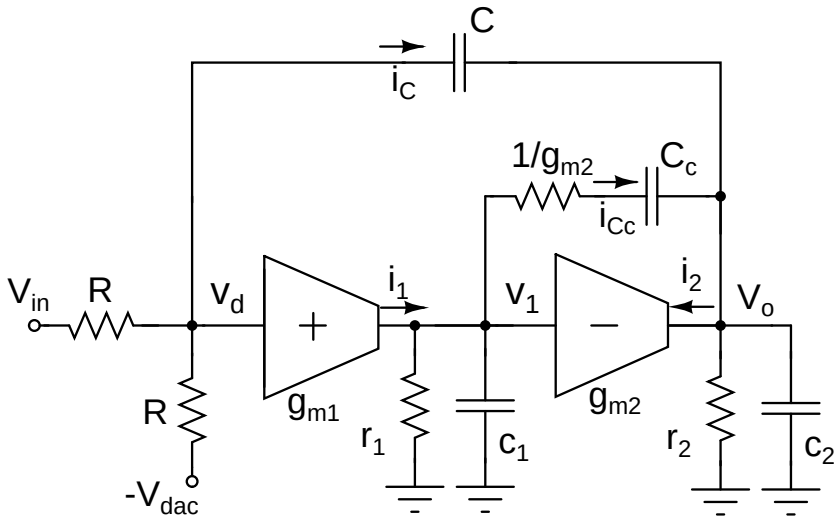


Figure 2.5: Circuit of the input stage of a CTDSM using an active-RC integrator with a two-stage Miller compensated opamp.

Analyzing nonlinearity in this integrator can be accomplished through the Volterra series. Unfortunately, this results in complex expressions that give little circuit intuition. We employ a simple, intuitive approach and apply Bussgang's method of current injection to model the nonlinearity as described below.

For simplicity, we begin with the assumption that the output conductances at each stage of the opamp is zero ($r_1 = r_2 = \infty$). With an NRZ DAC, the input to the integrator can be perceived as a series of step-like signals given by,

$$V_{in}(t) - V_{dac}(t) = \sum_n x[n]\,(u(t - nT_s) - u(t - (n+1)T_s)) \tag{2.20}$$

where $u(t)$ denotes the *Heaviside step function* and $x[n] = V_{in}(nT_s) - V_{dac}(nT_s)$, since the input signal can also be considered to be sampled and held on the grounds of oversampling. Note that a practical active-RC integrator (used as the first integrator in a CTDSM) is designed in a manner such that the parasitic poles lie much beyond the sampling frequency (f_s) used. In other words, it is fair to assume that the opamp is fast in a well-designed integrator. Understandably, the initial transients at the virtual ground node (v_d) can then be seen to quickly die in each clock period. This suggests that the response at v_d can also be considered as step signals, similar to the integrator input. Fig. 2.6 illustrates this scenario, where the responses expected at each node of the integrator is shown for a step input.

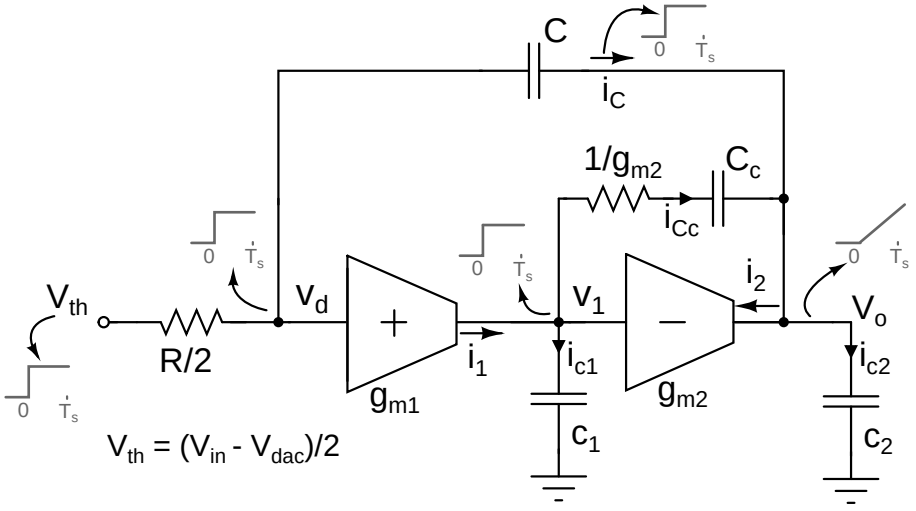


Figure 2.6: Circuit of the first integrator of a CTDSM with the responses (currents) shown at each node given a step input. Initial transients are neglected assuming a fast opamp.

With steps in the integrator input, the second stage transconductor (g_{m2}) is required to supply a constant current (i_2) in a given clock period. This implies that the opamp's first stage output voltage ($v_1 = i_2/g_{m2}$) is also a step signal in each clock cycle, thus resulting in $i_{c1} \approx 0$. The voltages across C and C_c being approximately the same, we therefore have

$$i_{Cc}\ (\approx i_1) \approx i_C \frac{C_c}{C} \qquad \text{or,} \tag{2.21}$$

$$i_C \approx \frac{C}{C_c} i_1 = \frac{C}{C_c} g_{m1} v_d \tag{2.22}$$

With $i_C = (V_{in} - V_{dac} - 2v_d)/R$, the above relation can be used to express the linear response at the virtual ground node as

$$v_{d,lin} = \frac{(V_{in} - V_{dac})}{2 + k g_{m1} R} \tag{2.23}$$

where $k = C/C_c$. It can be seen that the relation is similar to that in eq. 2.8, though the swing at the virtual ground node is smaller by a factor C/C_c in a Miller opamp based integrator. The knowledge of the linear response thus obtained can be used to inject a step-like current $i_{inj} = g_{31} v_{d,lin}^3$ as shown in Fig. 2.7.

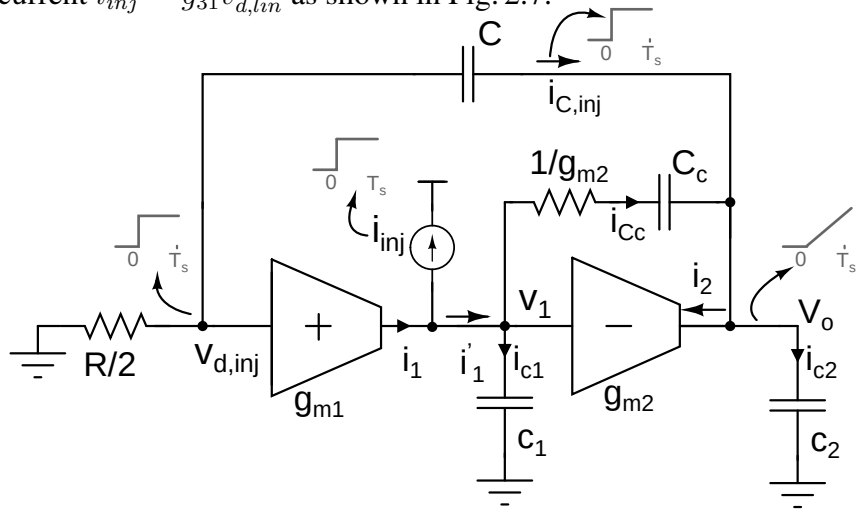


Figure 2.7: Injection of a step-like nonlinear current (i_{inj}) at the output of the input transconductor in an active-RC integrator with Miller opamp. Initial transients have been neglected.

The resulting step-like capacitive current $i_{C,inj}$ is given by

$$i_{C,inj} \approx ki_{Cc} \tag{2.24}$$

$$= k(i_1 - i_{inj}) = k(g_{m1}v_{d,inj} - i_{inj}) \tag{2.25}$$

With $i_{C,inj} = -2v_{d,inj}/R$, we have

$$\frac{-2v_{d,inj}}{R} \approx k(g_{m1}v_{d,inj} - i_{inj}) \quad \text{or,} \tag{2.26}$$

$$v_{d,inj} \approx \frac{kR}{(2 + kg_{m1}R)}i_{inj} = \frac{R}{(2 + kg_{m1}R)}(kg_{31}v_{d,lin}^3) \tag{2.27}$$

Comparing the expressions for $v_{d,lin}$ and $v_{d,inj}$ with that obtained with the single-stage integrator (eq. 2.8 and 2.9), it is seen that g_m and g_3 of the latter is simply replaced by kg_{m1} and kg_{31} with the two-stage opamp. Thus, similar to the single-stage case, the nonlinear two-stage integrator can be modeled with a nonlinear function ($f(x)$) preceding a linear integrator ($-\omega_1/s$), where

$$f(x) = x - \frac{2g_{31}}{g_{m1}(2 + kg_{m1}R)^3}x^3 \tag{2.28}$$

$$\omega_1 = \frac{kg_{m1}}{C(2 + kg_{m1}R)} \tag{2.29}$$

Comparing the nonlinear functions in eq. 2.18 and 2.28, an advantage of using Miller opamp can be brought about. With a two-stage implementation, an improvement in the integrator linearity can be obtained by increasing $k = C/C_c$. This is in contrast to the need for increasing either g_m or R, as in a single-stage implementation.

While the above analysis provides us with valuable insights about the circuit operation, very large output resistances cannot be guaranteed in all circumstances (especially r_2). Due to finite resistances, the factor k relating the capacitor current (i_C) and the

transconductor's current (i_1) gets modified in a manner as explained below.

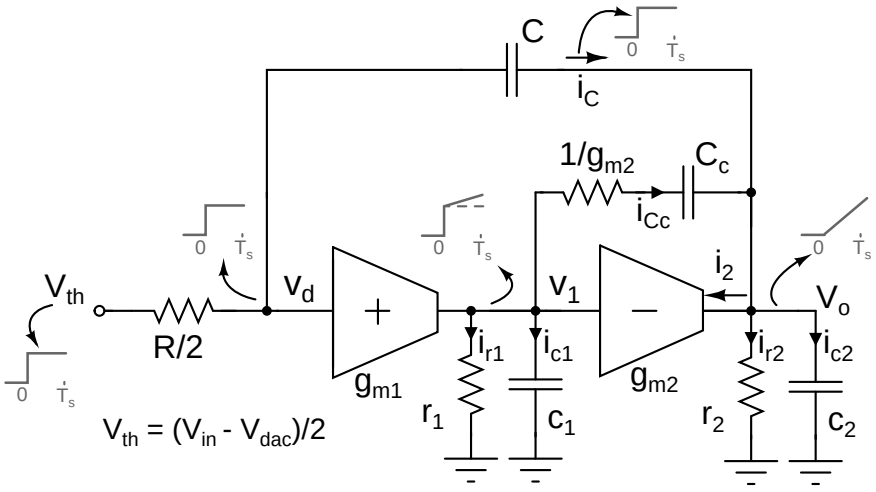


Figure 2.8: Circuit of the first integrator of a CTDSM with finite r_1 and r_2. The responses (currents) at each node are shown for a given step input, neglecting the initial transients.

As shown in Fig. 2.8, a step input to the integrator results in a step-like capacitive current (i_C), which gets integrated to result in a ramp at the output of the integrator. With a finite value for r_2 , this gives rise to a finite value for the current, $i_{r2} = V_o/r_2$. Since the second stage transconductor has to supply this current, the first stage output voltage can now be expressed as having two components given by,

$$v_1 = \frac{i_2}{g_{m2}} = v_{1,res} + v_{1,cap} \tag{2.30}$$

$$= \frac{-i_{r2}}{g_{m2}} + \frac{(i_C + i_{Cc} - i_{c2})}{g_{m2}} \tag{2.31}$$

where the component $v_{1,res}$ denotes the linearly increasing voltage with time and $v_{1,cap}$ represents the constant component required to supply the capacitive currents[4]. The time-varying component at v_1 can be observed to result in the capacitive currents (i_{Cc}

[4]These capacitive currents are present even when $r_2 = \infty$.

and i_{c1}) given by

$$i_{Cc} \approx sC_c(v_{1,res} - V_o) = -sC_cV_o(1 + \frac{1}{g_{m2}r_2}) \tag{2.32}$$

$$i_{c1} \approx sc_1v_{1,res} = -sc_1V_o\frac{1}{g_{m2}r_2} \tag{2.33}$$

With $i_C \approx -sCV_o$, we have

$$i_{Cc} \approx i_C\frac{C_c}{C}(1 + \frac{1}{g_{m2}r_2}) \tag{2.34}$$

$$i_{c1} \approx i_C\frac{c_1}{C}\frac{1}{g_{m2}r_2} \tag{2.35}$$

These currents are to provided by the first transconductor, g_{m1}. Additionally, with a finite value for r_1, the transconductor also needs to supply the resistive current given by

$$i_{r1} = \frac{v_1}{r_1} \approx \frac{(i_C + i_{Cc} - i_{c2})}{g_{m2}r_1} \tag{2.36}$$

$$\approx \frac{i_C(1 + C_c/C + c_2/C)}{g_{m2}r_1} \qquad (\text{from eq. } 2.34) \tag{2.37}$$

Combining the above equations, the first transconductor current (i_1) can be expressed as

$$i_1 = i_{Cc} + i_{c1} + i_{r1} \tag{2.38}$$

$$\approx i_C\left(\frac{C_c}{C}(1 + \frac{1}{g_{m2}r_2}) + \frac{c_1}{C}\frac{1}{g_{m2}r_2} + \frac{1 + C_c/C + c_2/C}{g_{m2}r_1}\right) \tag{2.39}$$

Grouping the terms, we thus obtain

$$k = \frac{i_C}{i_1} \approx \frac{C}{C_c + \frac{c_1 + C_c}{g_{m2}r_2} + \frac{C + C_c + c_2}{g_{m2}r_1}} \tag{2.40}$$

which approximates to C/C_c for high values of r_1 and r_2, as expected. The expression thus derived was confirmed through simulations. Fig. 2.9 shows the simulated step response at the virtual ground node (of a linear integrator) along with the expected response as predicted using the gain factor in the relation

$$v_{d,lin} = \frac{(V_{in} - V_{dac})}{2 + kg_{m1}R} \tag{2.41}$$

with k being determined from eq. 2.40. For comparison purposes, the response predicted by assuming $k = C/C_c$ is also shown. It can be seen that the predicted responses closely match the simulated ones[5]. Understandably, better accuracy achieved with k obtained from eq. 2.40 is more apparent for r_2 = 200 kΩ.

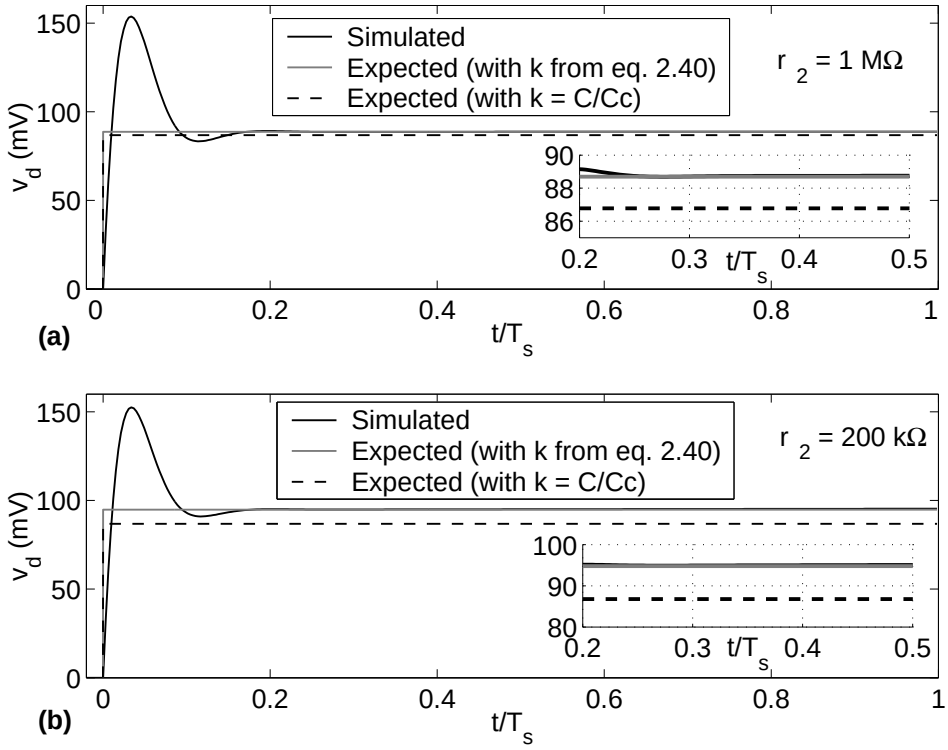


Figure 2.9: Simulated and expected step response at v_d for the Miller opamp based integrator with g_{m1} = 27.12 μS, R = 100 kΩ, C = 527 fF, C_c =150 fF, g_{m2} = 60 μS for (a) r_2 = 1 MΩ (b) r_2 = 200 kΩ. The insets show the zoomed-in views.

[5]Note that we have neglected the initial transients in this analysis.

Thus, using the expression for $k = i_C/i_1$ (as given in eq. 2.40), the nonlinear integrator can be modeled with

$$f(x) = x - \frac{2g_{31}}{g_{m1}(2 + kg_{m1}R)^3}x^3 \quad (2.42)$$

$$\omega_1 = \frac{kg_{m1}}{C(2 + kg_{m1}R)} \quad (2.43)$$

In summary, we have modeled the nonlinearity in a Miller based active-RC integrator by using a simple, intuitive approach without resorting to complex analysis involving Volterra series. Neglecting the initial transients generated due to the step inputs $(V_{in} - V_{dac})$ applied to the integrator, we related the integrating capacitor's current (i_C) and the first transconductor's current (i_1). This enabled us to easily model the nonlinear integrator, in a manner similar to that of a single-stage integrator. An important observation from the analysis is the possibility to control the integrator nonlinearity through the factor $k \approx C/C_c$. This can be achieved by either increasing the value of C or by decreasing C_c. However, any reduction in the value of C_c will compromise the stability of the integrator. We shall now investigate an integrator topology that can effect this reduction without compromising on the stability.

3) Active-RC integrator with feedforward compensated opamp: Consider an active-RC integrator built with two-stage feedforward compensated opamp as shown in Fig. 2.10. The feedforward transconductor (g_{m3}) serves to bring stability to the closed loop system. An advantage of this compensation scheme is that the bandwidth (or speed) of the opamp is not reduced, unlike with Miller compensation. The parasitic Miller capacitance of the second stage transconductor, g_{m2} has been modeled as c_p. This can be used for comparing the nonlinearity with that observed in the case of a Miller compensated opamp.

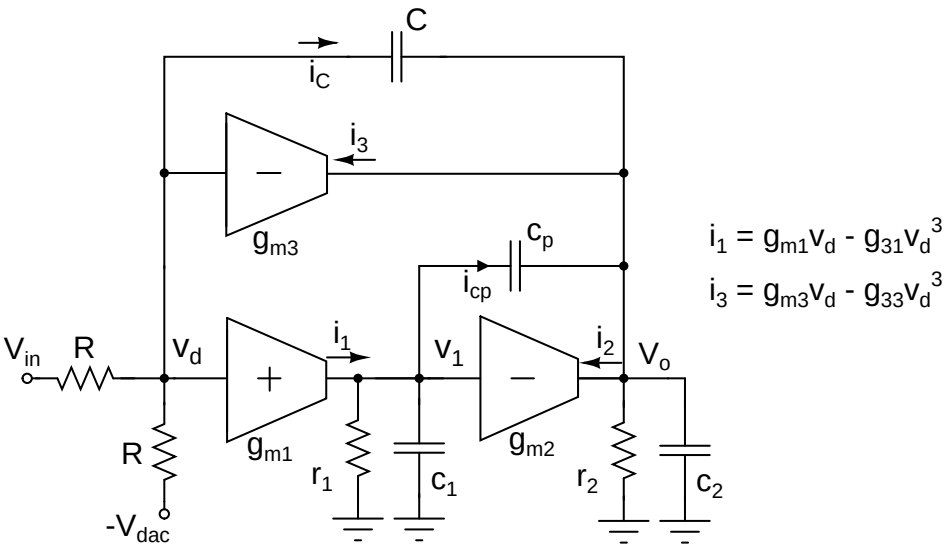


Figure 2.10: Circuit of an active-RC integrator with a two-stage feedforward compensated opamp.

Since significant part of this integrator topology is similar to that with a Miller opamp based integrator, the same analysis can be carried over here. With a linear integrator, in response to a step input, the presence of the feedforward transconductor (g_{m3}) will modify the step-like voltage component at v_1 as

$$v_{1,cap} = \frac{(i_C + i_{cp} - i_{c2} - i_3)}{g_{m2}} \tag{2.44}$$

where $i_3 = g_{m3}v_d$ denotes the current from the feedforward transconductor. This results in the modified resistive current given by

$$i_{r1} \approx \frac{v_{1,cap}}{r_1} \tag{2.45}$$

$$= \frac{(i_C + i_{cp} - i_{c2})}{g_{m2}r_1} - \frac{i_3}{g_{m2}r_1} \tag{2.46}$$

Rewriting the expression for i_1,

$$i_1 = i_{cp} + i_{c1} + i_{r1} \tag{2.47}$$

$$\approx i_C \left(\frac{c_p}{C}(1 + \frac{1}{g_{m2}r_2}) + \frac{c_1}{C}\frac{1}{g_{m2}r_2} + \frac{1 + c_p/C + c_2/C}{g_{m2}r_1} \right) - \frac{i_3}{g_{m2}r_1} \tag{2.48}$$

Denoting $k = \dfrac{C}{c_p + \dfrac{c_1 + c_p}{g_{m2}r_2} + \dfrac{C + c_p + c_2}{g_{m2}r_1}}$, we thus have

$$i_C = ki_1 + \frac{k}{g_{m2}r_1}i_3 \tag{2.49}$$

$$= k\left(g_{m1} + \frac{g_{m3}}{g_{m2}r_1}\right)v_d \tag{2.50}$$

Comparing the above relation with eq. 2.22, the relevant expressions for the feedforward opamp based integrator can be simply obtained by replacing g_{m1} in the expressions for the Miller opamp case with $g_{m1} + g_{m3}/(g_{m2}r_1)$. The component due to g_{m3} can be considered as the current from the feedforward transconductor, as referred to the the first stage output of the opamp using the factor '$g_{m2}r_1$'. Defining the nonlinear currents to be $i_1 = g_{m1}v_d - g_{31}v_d^3$ and $i_3 = g_{m3}v_d - g_{33}v_d^3$, the nonlinear integrator can thus be modeled with

$$f(x) \approx x - \frac{2g_{31}}{g_{m1}(2 + kg_{m1}R)^3}x^3 - \frac{2g_{33}/(g_{m2}r_1)}{g_{m1}(2 + kg_{m1}R)^3}x^3 \tag{2.51}$$

$$\omega_1 \approx \frac{kg_{m1}}{C(2 + kg_{m1}R)} \tag{2.52}$$

where it is assumed that $g_{m1} \gg g_{m3}/(g_{m2}r_1)$. Two important conclusions can be drawn from the above analysis - one, the nonlinearity of the feedforward transconductor should not be as critical as that of the input transconductor due to the scaled down nature of its current. Secondly, the expression for k indicates that the effect of the input transconductor's nonlinearity will be less pronounced with a feedforward compensated opamp, since c_p will always be smaller than C_c used in a Miller opamp. Physically, such an improvement can be attributed to the fact that the input transconductor does not have to charge/discharge any compensating capacitor C_c, in case of the feedforward opamp.

With reduced current requirements on the input transconductor, it can be seen that a feedforward opamp based active-RC integrator will be more linear than its Miller counterpart, for the same power consumption[6].

It has to be pointed out to the reader that an interesting case arises when the output resistances of the feedforward opamp are high and the parasitic Miller capacitance, $c_p \approx 0$. With such an opamp, eq. 2.51 predicts that the integrator is linear (with $k = \infty$), irrespective of the nonlinearity of the input transconductor. However, this will not be the case, since the initial transients at the virtual ground node[7] will reflect the nonlinearity in such integrators. Further, simulations show that the improvement expected with a feedforward opamp (over the Miller opamp) is slightly reduced due to the presence of these initial transients. It is thus seen that the initial transients should also be included while modeling the nonlinearity of such integrators for accurate characterizations. However, for reasons of brevity, we defer the modeling of the effect of initial transients till the next chapter.

To summarize the modeling discussed so far, we have shown that a nonlinear integrator can be modeled as a linear integrator preceded by a nonlinear function, $f(x)$. Starting with a simple G_m-C integrator, the model has been obtained for active-RC integrators with different kinds of opamp architectures. We shall now proceed to see how these models can be used to determine the performance of multi-bit CTDSMs.

[6]The reader can refer to Appendix B for a detailed comparison between the two opamp architectures.
[7]neglected in this analysis

2.2 Effect of integrator nonlinearity in multi-bit CIFF modulators

Based on the discussion in the previous section, the nonlinearity of the first integrator in a CIFF CTDSM can be modeled as shown in Fig. 2.11, where the function, $f(x) = x - \beta x^3$ models the weak nonlinearity in the integrator. The zeroth-order-hold (ZOH) in the feedback path can model the NRZ nature of the feedback pulse shape. Using the assumptions stated below, we will see how the effect of integrator nonlinearity can be quantified with this model.

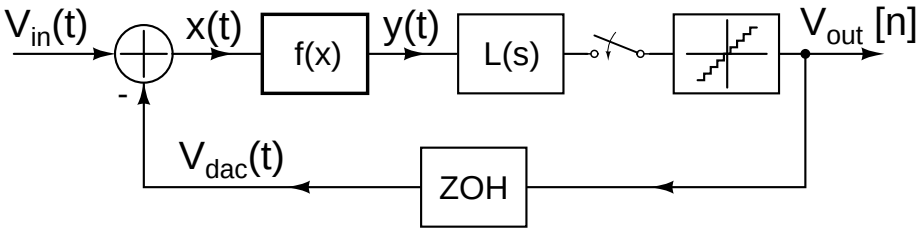


Figure 2.11: CTDSM with nonlinearity in the first integrator

Assumptions behind the analysis

1. *Quantization noise is white and uncorrelated with the input:* This is the usual assumption, which is largely satisfactory in multi-bit quantizers.

2. *Input signal is constant within a clock period:* As assumed in the previous section, if the input is slowly varying, it is reasonable to approximate it by a waveform which is piecewise constant. This simplifies the analysis of nonlinearity in the modulator and also allows us to express the input to the first integrator as a sampled and held version of $x[n] = V_{in}[n] - V_{out}[n]$.

3. *In-band Signal Transfer Function (STF) is 1:* If the dc gain of the loop filter is sufficiently high, the in-band STF is close to unity, implying that the input to the first integrator ($x[n]$) is the shaped quantization noise. In reality, there is a small signal component, which we neglect in this analysis.

4. *Shaped quantization noise sequence is jointly Gaussian:* Denoting the quantization noise and the impulse response corresponding to the NTF by $e[n]$ and $h[n]$ respectively, we see that the shaped noise sequence is given by $e[n] * h[n]$ (where $*$ is the convolution operation). Since the shaped noise is a filtered version of $e[n]$, it can be considered as a running sum of independent, delayed samples of $e[n]$. By Central Limit Theorem, this can be assumed to result in a Gaussianly distributed random process.

To validate our assumptions before analysis, simulations were run[8] to observe the distribution of the input to the first integrator, $x[n]$. Fig. 2.12 shows the histogram observed with a sinusoidal input for different number of quantizer levels[9]. The expected distribution of the shaped quantization noise is also shown for comparison, assuming uniform distribution for $e[n]$. It can be seen that there is good accordance between the observed and expected histograms. Also, the observed distribution tends closer and closer to the predicted one, with increasing number of quantizer levels. This is because of the uniform assumption becoming more and more valid for increasing quantizer levels. The effect of varying the OBG was also observed with a 4-bit modulator, as shown in Fig. 2.13 and 2.14. The histograms follow the predicted distribution more closely for higher OBG modulators, due to the fact that aggressive NTFs make the quantization noise more random in nature.

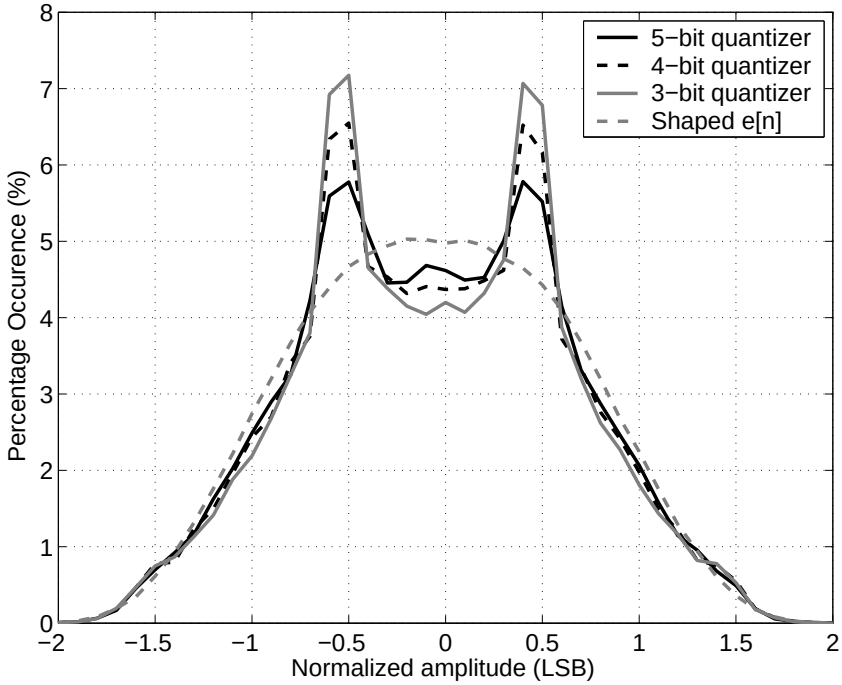


Figure 2.12: Observed histogram of samples of $x[n]$ for various quantizer levels. Expected distribution of noise shaped $e[n]$ is also shown for comparison.

[8]The modulator was assumed to have a maximally flat NTF (OBG = 3) with complex in-band zeros.

[9]The peaks observed in the distribution around ± 0.5 LSB are because of the amplitude of the input sinusoid. Depending on the difference between the amplitude and the nearest quantizer level, the peaks can be expected to occur in the range of [-0.5 0.5]LSB

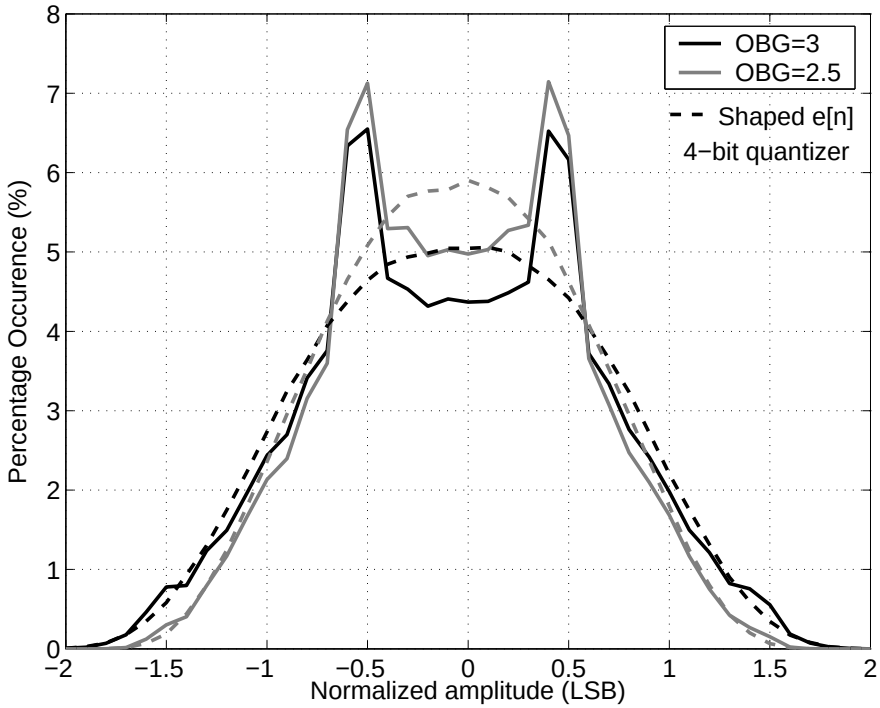


Figure 2.13: Observed histogram of samples of $x[n]$ for various OBGs. The expected distribution is shown as dashed-lines.

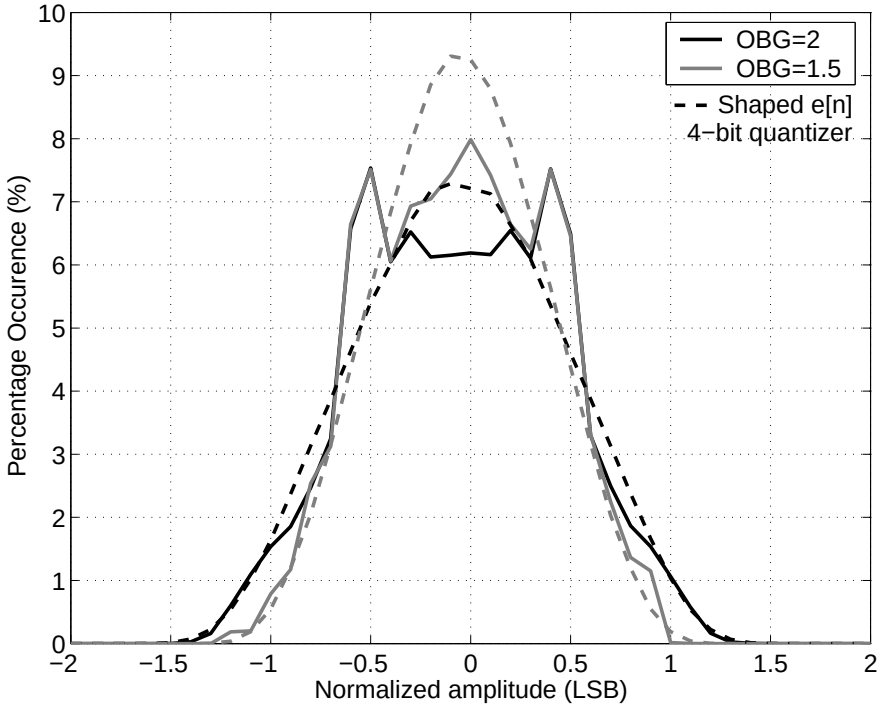


Figure 2.14: Observed histogram of samples of $x[n]$ for various OBGs. The expected distribution is shown as dashed-lines.

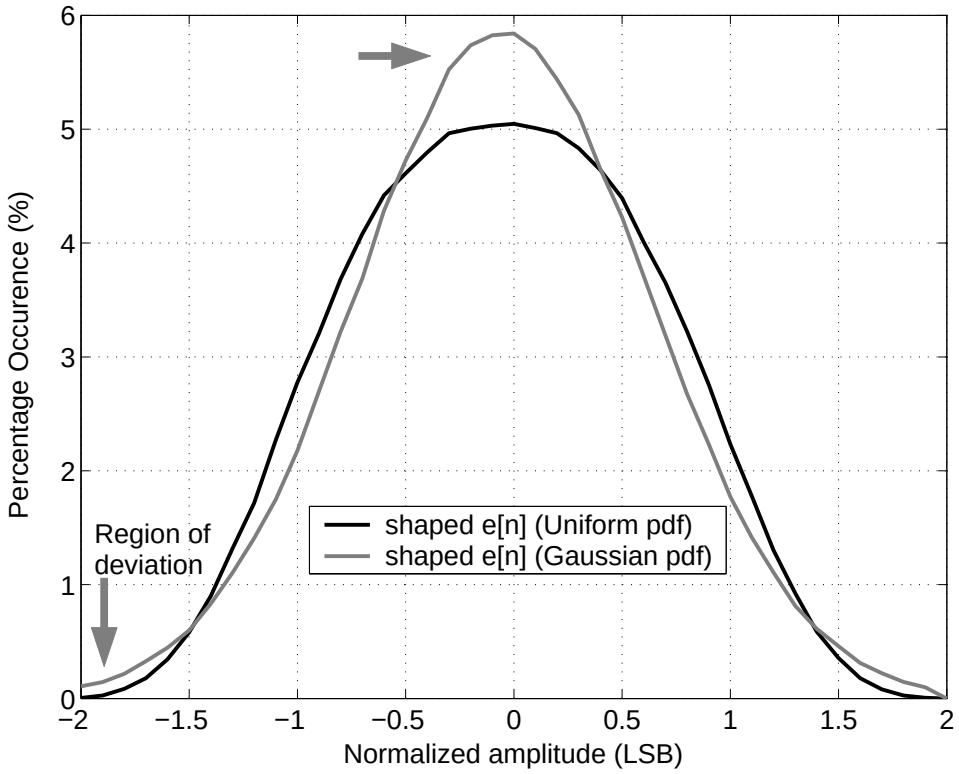


Figure 2.15: Comparison of the expected distribution of noise shaped $e[n]$ having either uniform or Gaussian distribution. The regions where the two distributions differ have been indicated.

While the veracity of white noise and uniform assumption for $e[n]$ can be understood from the distributions shown thus far, Fig. 2.15 can provide further insights on the Gaussian assumption. For comparison purposes, the distribution of the shaped quantization noise was also obtained by assuming a Gaussian distribution for $e[n]$. Comparison between the two distributions clearly shows slight differences in their densities, observed along the extremes and around the mean value. As a result, the Gaussian assumption for the shaped quantization noise in a CTDSM can in principle be expected to lead to errors. Nevertheless, as can be seen from the discussion below, such an assumption allows the use of classical results on the effect of nonlinear functions of Gaussian processes. This enables us to quantify the effect of integrator nonlinearity in CTDSMs. Further, simulations show that the error incurred is only marginal and the assumption is

reasonable in practical multi-bit modulators.

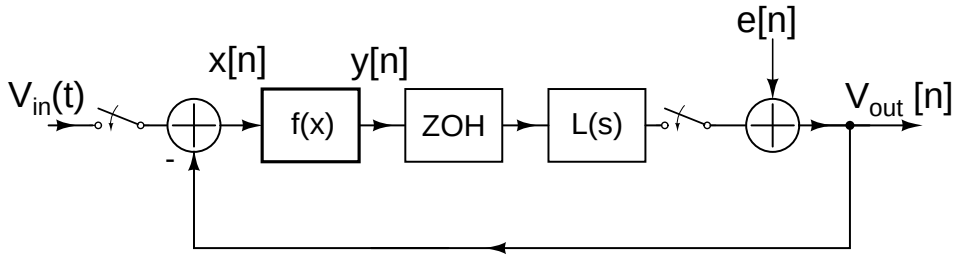


Figure 2.16: Equivalent model of the CTDSM for slow inputs

Using the above mentioned assumptions on the input signal and $e[n]$, the CTDSM can be equivalently modeled as shown in Fig. 2.16. The quantization noise has been modeled as an additive error and the ZOH block has been brought into the forward path[10] as it is now common to both V_{in} and V_{out}. For simplicity, we first consider the case where $V_{in}(t) = 0$. Then, the input to the loop filter ($x[n]$) is the shaped quantization noise, with an autocorrelation, $R_{xx}[n] = (\Delta^2/12)h[n] * h[-n]$. Assuming $x[n]$ to be a Gaussian process, the autocorrelation function of the sequence, $y[n]$ observed at the output of the nonlinearity $f(x) = x - \beta x^3$ can be shown to be (see Appendix C)

$$R_{yy}[n] \approx R_{xx}[n] + 6\beta^2 R_{xx}^3[n] \tag{2.53}$$

From the above equation, the effect of integrator nonlinearity can be thought of as adding a random sequence, $z[n]$ to the input of an otherwise linear modulator, as shown in Fig. 2.17.

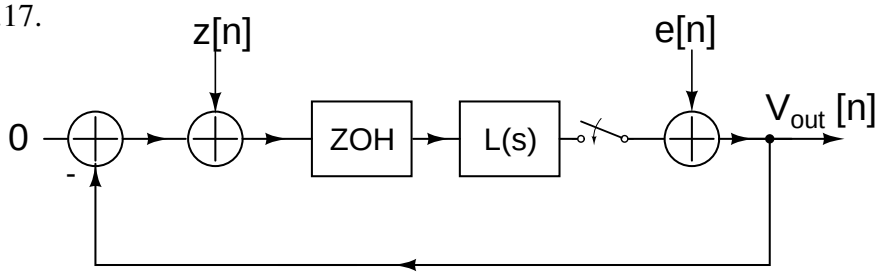


Figure 2.17: Model for nonlinearity analysis of the CTDSM for the zero input case.

[10]Note that the nonlinear function and the ZOH have been swapped, without loss of accuracy.

From eq. 2.53, the autocorrelation function of the random sequence is given by

$$R_{zz}[n] \;=\; 6\beta^2 R_{xx}^3[n] \tag{2.54}$$

In terms of the power spectral densities[11], we thus have

$$S_{yy}(\omega) \approx S_{xx}(\omega) + S_{zz}(\omega) \tag{2.55}$$

where $S_{zz}(\omega)$ is the Fourier transform of $R_{zz}[n]$. With the in-band STF being unity, the output of the modulator thus consists of an in-band noise which is composed of the usual shaped quantization noise along with $S_{zz}(\omega)$ (arising due to integrator nonlinearity). The PSD of $S_{zz}(\omega)$ can be easily determined once the modulator NTF and the details of the nonlinearity (β) are known, thus enabling us to quantify the effect of integrator nonlinearity in multi-bit CTDSMs. Physically, this IBN can be perceived as arising from the "demodulation" of various out-of-band tones due to the nonlinearity, $f(x)$. With maximally flat NTFs, this demodulation is found to result in a noise, which is flat within the signal band. This explains the variation in the in-band noise floor observed in our preliminary simulations (in Chapter 1). To gain more quantitative insights on the influence of the NTF on the in-band noise due to nonlinearity, a rough estimate of the IBN spectral density can be obtained by interpreting eq. 2.55 as discussed below.

Simplified expression for IBN and discussion:

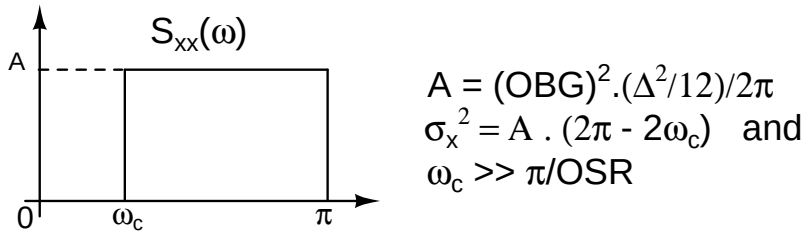


Figure 2.18: Assumed PSD of the shaped quantization noise

[11]The frequency variable ω refers to the discrete-time angular frequency expressed in rad/sample.

Assuming maximally flat NTFs, the PSD of the shaped quantization noise ($S_{xx}(\omega)$) can be approximated by a brick-wall function as shown in Fig. 2.18. With a given OBG and quantizer step size (Δ), $S_{zz}(\omega)$ can be obtained by circular convolution of $S_{xx}(\omega)$. Since the in-band PSD is observed to be flat, the increase in the in-band noise can be determined by evaluating $S_{zz}(\omega)$ for $\omega \approx 0$ as

$$S_{zz}(0) \;=\; 6\beta^2 A^3 \left[4(\pi \,-\omega_c)^2 \,-\, (\pi^2 - \omega_c^2)\right] \tag{2.56}$$

The above equation clearly proves our initial observation that increasing OBG results in an increased IBN ($\propto OBG^6$). Note that this leads to a similar kind of a trade-off as with the issues of clock jitter and time constant variations for the following reason - for smaller values of OBG, when the IBN is dominated by shaped quantization noise, any increase in OBG results in an equivalent decrease in the in-band noise. However, at higher values of OBG, the IBN is dominated by the noise due to nonlinearity, eventually causing it to increase as OBG is raised. It is thus seen that there is an optimal OBG that results in the least IBN. Understandably, this optimum depends on the amount of nonlinearity, β. Eq. 2.56 also indicates that decreasing the resolution of the quantizer by one bit causes $S_{zz}(0)$ (and the IBN due to nonlinearity) to increase by 18 dB. This is exactly the same performance difference seen between a 3-bit and 4-bit modulator, in our preliminary simulations (see Fig. 1.31). This also confirms the general intuition that increasing the number of quantizer levels greatly reduces the effect of nonlinearity.

We derived eq. 2.55 based on a zero input into the modulator. However, even when the input is non-zero, the input to the first integrator is practically the shaped quantization noise (since the STF is unity). Thus, it seems reasonable that the results that were

derived for the zero input case are valid even for a non-zero input. However, as will be shown in the following discussion, this is not true for modulators with RZ DAC.

2.2.1 Analysis of modulators with an RZ DAC

The effect of integrator nonlinearity in multi-bit CTDSMs with an RZ DAC can be analyzed similar to the case with an NRZ DAC. Increased tolerance to excess loop delay (upto half a clock cycle) and mitigation of the effects of asymmetry in the rise and fall times of the DAC, are reasons for preferring RZ DAC over an NRZ DAC. A multi-bit CTDSM with an RZ DAC and a nonlinear first integrator is shown in Fig. 2.19, where the RZ DAC has been depicted in the feedback path along with its impulse-response.

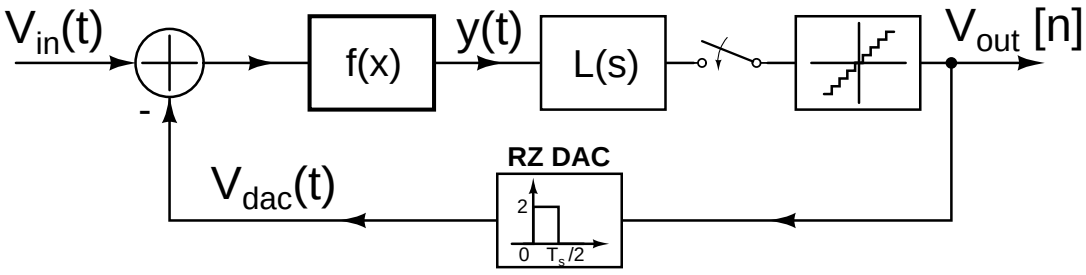


Figure 2.19: CTDSM employing RZ DAC with nonlinearity in the first integrator

While the assumptions made with the NRZ DAC hold good for an RZ DAC, observe that the input to the first integrator ($y(t)$) or the nonlinear function is not simply a sampled and held version of $V_{in}[n] - V_{out}[n]$. This necessitates a modified approach to analyze the effect of integrator nonlinearity and we therefore develop the theory by considering the cases with and without input signal, separately.

A. When modulator input is zero : Assuming zero input to the modulator allows us to consider the input to the nonlinear function as the shaped quantization noise sequence

$(x[n])$ scaled by the RZ DAC. Therefore, the integrator input is given by

$$y(t) = 2x[n] - 8\beta x^3[n], \quad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+0.5)\mathrm{T_s} \tag{2.57}$$

$$= 0, \qquad (\mathrm{n}+0.5)\mathrm{T_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s} \tag{2.58}$$

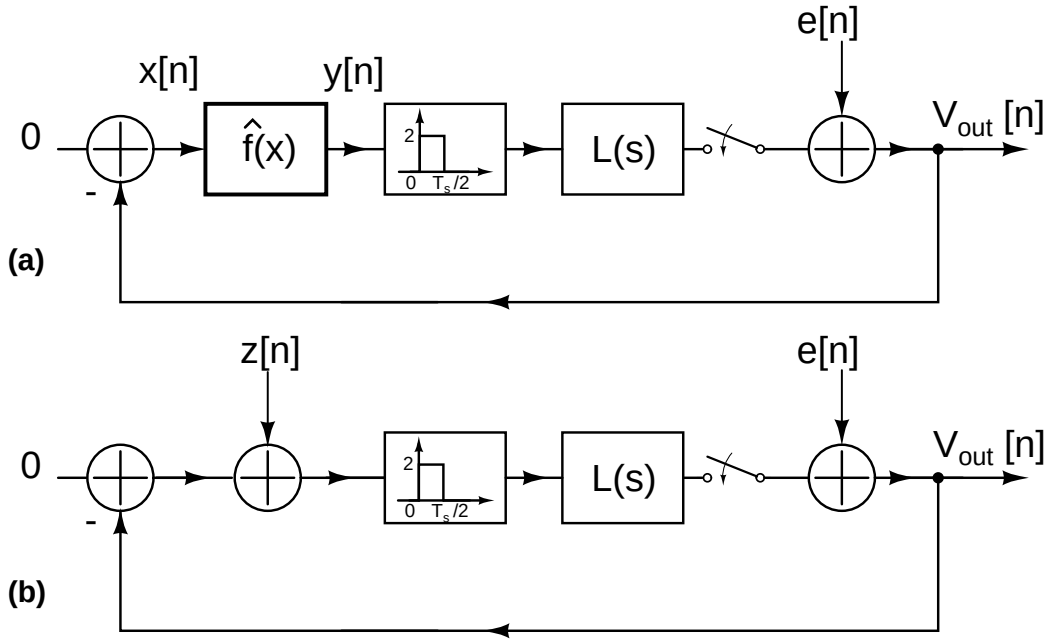


Figure 2.20: (a) CTDSM employing RZ DAC with zero input (b) Equivalent model for the CTDSM with a nonlinear first integrator, for zero inputs.

As shown in in Fig. 2.20(a), the RZ DAC can now be equivalently pushed in to the forward path resulting in a modified nonlinear function $(\widehat{f}(x))$ given by

$$\widehat{f}(x) = x - 4\beta x^3 \tag{2.59}$$

Similar to the NRZ case, integrator nonlinearity can thus be perceived to introduce the random sequence, $z[n]$ as a separate input to the linear modulator depicted in Fig. 2.20(b). Using the Gaussian assumption for $x[n]$, the autocorrelation function of this random sequence can be found as

$$R_{zz}[n] = 96\beta^2 R_{xx}^3[n] \tag{2.60}$$

The increase in the in-band noise due to integrator nonlinearity can therefore be quantified by interpreting the above equation in frequency domain, as before. Comparing eq. 2.54 and 2.60, it can be seen that the IBN due to nonlinearity is 12 dB worse with an RZ DAC when compared to an NRZ implementation.

B. When modulator input is non-zero : The scenario with non-zero input can be analyzed by representing the input signal path as shown in Fig. 2.21(a). The input signal has been assumed to be sampled and held through two RZ DACs - with one of them being delayed by half a clock cycle. Such a representation permits us to have an RZ DAC in the forward path, common to both $V_{in}[n]/2$ and $V_{out}[n]$, as in Fig. 2.21(a).

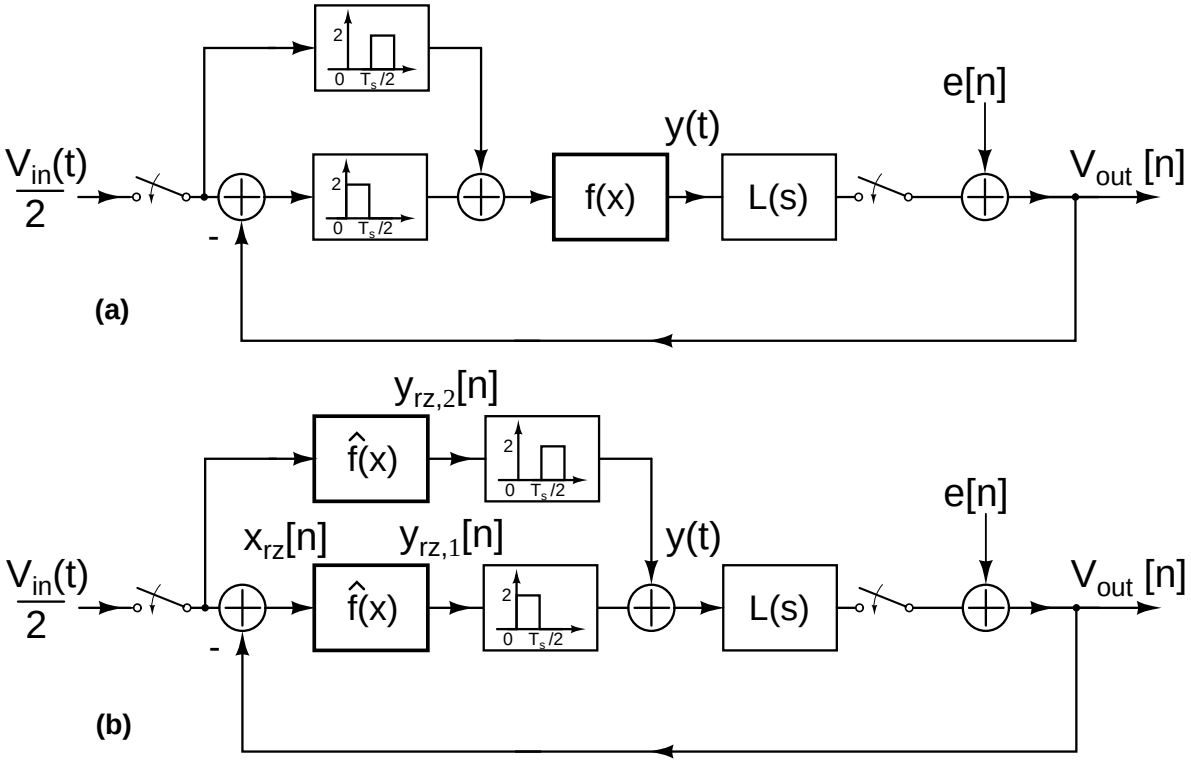


Figure 2.21: (a) Model for the CTDSM employing RZ DAC with non zero input and (b)its equivalent representation.

Since the respective outputs of the RZ DACs are interleaved in time, without loss of accuracy, the nonlinear function ($f(x)$) can be brought into the respective signal paths (as shown in Fig. 2.21(b)). The sequence $x_{rz}[n]$, in the path involving $V_{out}[n]$, can now

be expressed in terms of the shaped quantization noise ($x[n] = V_{in}[n] - V_{out}[n]$) as

$$x_{rz}[n] = \frac{1}{2}V_{in}[n] - V_{out}[n] \tag{2.61}$$

$$= x[n] - \frac{1}{2}V_{in}[n] \tag{2.62}$$

With such a representation, evaluating the output of the nonlinear function results in

$$y_{rz,1}[n] = x[n] - \frac{1}{2}V_{in}[n] - 4\beta\left(x[n] - \frac{1}{2}V_{in}[n]\right)^3 \tag{2.63}$$

$$= x[n] - \frac{1}{2}V_{in}[n] - 4\beta x^3[n] + \frac{\beta}{2}V_{in}^3[n] - 3\beta x[n]V_{in}^2[n] + 6\beta x^2[n]V_{in}[n] \tag{2.64}$$

$$y_{rz,2}[n] = \frac{1}{2}V_{in}[n] - \frac{\beta}{2}V_{in}^3[n] \tag{2.65}$$

Several observations can be drawn from the above equations.

- From eq. 2.64, observe that the first of the nonlinear terms ($-4\beta x^3[n]$) is the same component derived with zero input.
- While the second nonlinear term involving the cube of the input signal has the potential to lead to third harmonic, such tones in $y_{rz,1}(t)$ and $y_{rz,2}(t)$ will get combined to result in $y(t)$ given by

$$y(t) = \frac{\beta}{2}V_{in}^3[n]\,, \qquad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+0.5)\mathrm{T_s} \tag{2.66}$$

$$= -\frac{\beta}{2}V_{in}^3[n]\,, \quad (\mathrm{n}+0.5)\mathrm{T_s} \leq \mathrm{t} < \mathrm{nT_s} \tag{2.67}$$

  It can be seen that the harmonics generated get modulated out of the signal band[12].
- It can be seen that the noise due to the last two terms in eq. 2.64 is dependent on the input signal, V_{in}. The squaring operation on $x[n]$ in the term $6\beta x^2[n]V_{in}[n]$ can result in the demodulation of out-of-band noise into the in-band region. This can lead to significant increase in the in-band noise of the modulator. Since there is no such demodulation process occurring with the component $3\beta x[n]V_{in}^2[n]$, it can be neglected in the analysis for all practical purposes.

Thus, simplifying the expressions for $y_{rz,1}[n]$ and $y_{rz,2}[n]$ to have only the significant

[12]The anti-aliasing property of the loop filter can be assumed to attenuate the harmonics generated about $\pm f_s, \pm 3f_s, \cdots$

components, we have

$$y_{rz,1}[n] \approx x[n] - \frac{1}{2}V_{in}[n] - 4\beta x^3[n] + 6\beta x^2[n]V_{in}[n] \quad (2.68)$$

$$y_{rz,2}[n] \approx \frac{1}{2}V_{in}[n] \quad (2.69)$$

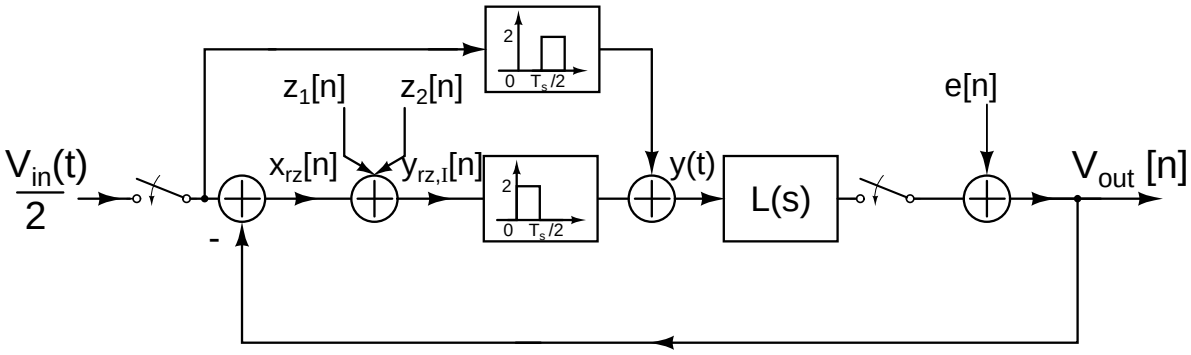


Figure 2.22: Equivalent model for the CTDSM employing RZ DAC with a nonlinear first integrator

The modulator with a nonlinear first integrator can thus be equivalently modeled as a linear modulator with two additional random inputs, $z_1[n] = -4\beta x^3[n]$ and $z_2[n] = 6\beta x^2[n]V_{in}[n]$, as shown in Fig. 2.22. Assuming $z[n] = z_1[n] + z_2[n]$ and A_{in} to denote the input signal amplitude, the PSD of the random sequence arising due to nonlinearity can be given by (see Appendix C)

$$S_{zz}(\omega) = S_{z_1z_1}(\omega) + S_{z_2z_2}(\omega) + 2Re\left(S_{z_1z_2}(\omega)\right) \quad (2.70)$$

$$= 96\beta^2\,\mathcal{F}\left[R_{xx}^3[n]\right] + 36\beta^2 A_{in}^2\,\mathcal{F}\left[R_{xx}^2[n]\right] + 2Re\left(S_{z_1z_2}(\omega)\right) \quad (2.71)$$

where $S_{z_1z_1}(\omega)$ and $S_{z_2z_2}(\omega)$ are the Fourier transforms of the respective auto-correlation functions while $S_{z_1z_2}(\omega)$ denotes the cross power spectral density component. A quick interpretation of the above equation using the brick-wall shaped PSD for $S_{xx}(\omega)$ (Fig. 2.18) can tell us the relative significance of the individual components. From the previous section, the flat PSD of $S_{z_1z_1}(\omega)$ in the in-band region can be approximated to

$$S_{z_1z_1}(0) = 96\beta^2A^3\left[4(\pi - \omega_c)^2 - (\pi^2 - \omega_c^2)\right] \quad (2.72)$$

where $A = OBG^2(\Delta^2/12)/2\pi$. Correspondingly, the PSD of $S_{z_2z_2}(\omega)$ can also be determined through circular convolution of the brick-wall PSD for $S_{xx}(\omega)$. The in-band noise spectral density thus obtained can be given by

$$S_{z_2z_2}(0) = 36\beta^2 A_{in}^2 A^2 (2\pi - 2\omega_c) \tag{2.73}$$

For simplicity, the worst case noise densities of $S_{z_1z_1}(0)$ and $S_{z_2z_2}(0)$ can be compared by assuming $\omega_c \approx 0$. Thus,

$$\frac{S_{z_2z_2}(0)}{S_{z_1z_1}(0)} \approx \frac{A_{in}^2}{4\pi A} \tag{2.74}$$

$$= \frac{A_{in}^2}{2OBG^2(\Delta^2/12)} \tag{2.75}$$

With a full-scale range of $2V_{ref}$ for a quantizer that has 2^{N} levels, $\Delta = 2V_{ref}/2^N$. Assuming a hypothetical peak signal amplitude of V_{ref} results in

$$\frac{S_{z_2z_2}(0)}{S_{z_1z_1}(0)} \approx \frac{3}{2}\left(\frac{2^N}{OBG}\right)^2 \tag{2.76}$$

For multi-bit quantizers, since $2^N > OBG$, $S_{z_2z_2}(\omega)$ can be observed to be more significant than $S_{z_1z_1}(\omega)$. Therefore, for peak input signals, $S_{zz}(\omega) \approx S_{z_2z_2}(\omega)$, thus significantly degrading the peak SNR and the effective resolution achieved with an RZ modulator. It was earlier shown that the performance degradation with an RZ DAC is 12 dB worse than with an NRZ DAC, when only $S_{z_1z_1}(\omega)$ is considered. This scenario is bound to worsen further in the presence of input signal, when the in-band noise is primarily determined by $S_{z_2z_2}(\omega)$. In these cases, the (peak) IBN due to nonlinearity can be seen to increase by 12 dB when the quantizer resolution is decreased by one bit.

2.3 Simulation results

A third order CTDSM with a 4-bit quantizer and a maximally flat NTF was used as a test vehicle to verify the results derived in the previous section. The sampling frequency and signal bandwidth were chosen to be 3.072 MHz and 24 kHz, respectively. A single-stage active-RC integrator (as in Fig. 2.3(a)) with R = 100 kΩ was used as the first integrator. All the other blocks of the modulator were assumed to be ideal. Throughout this section, unless specified otherwise, the feedback DAC is of the NRZ kind.

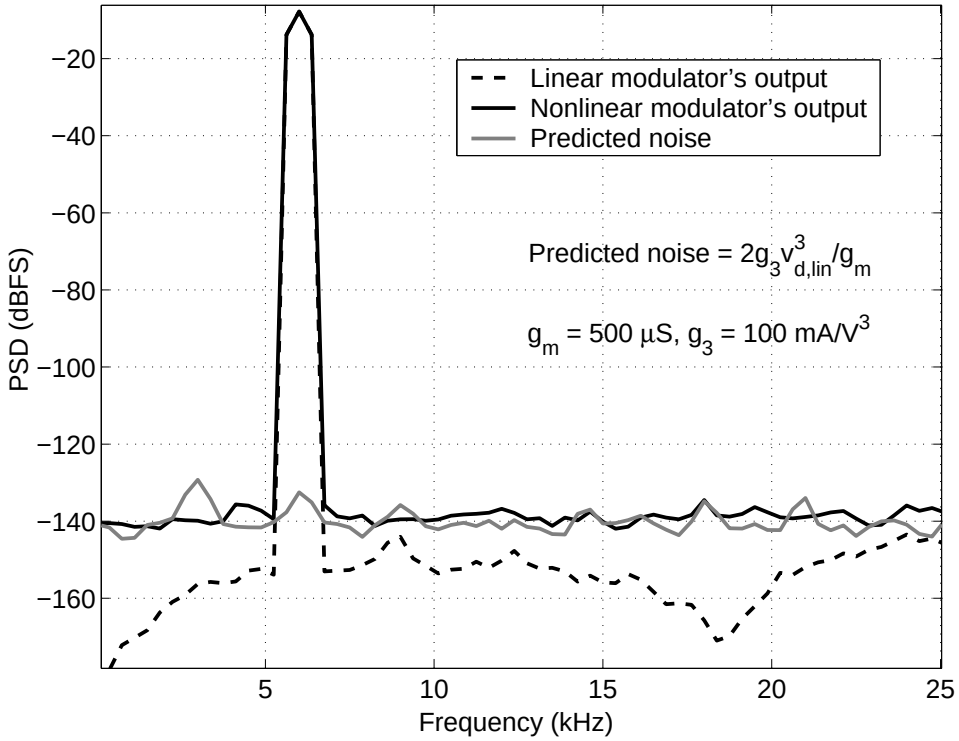


Figure 2.23: In-band power spectral density of the modulator output with a linear and nonlinear active-RC integrator. Predicted spectral density obtained using current injection method is also shown.

As a first step, the accuracy of the current injection method was confirmed by simulating a 4-bit modulator, assuming g_m = 500 μS and g_3 = 100 mA/V^3 for the transconductor. Observe from eq. 2.18 that the expected increase in the in-band noise floor can be determined using $2g_3v_{d,lin}^3/g_m$, where $v_{d,lin} = (V_{in} - V_{dac})/(2 + g_mR)$ is obtained by

simulating the linear modulator. Fig. 2.23 shows how well the predicted noise matches with the simulated spectral density, thus verifying the method of current injection.

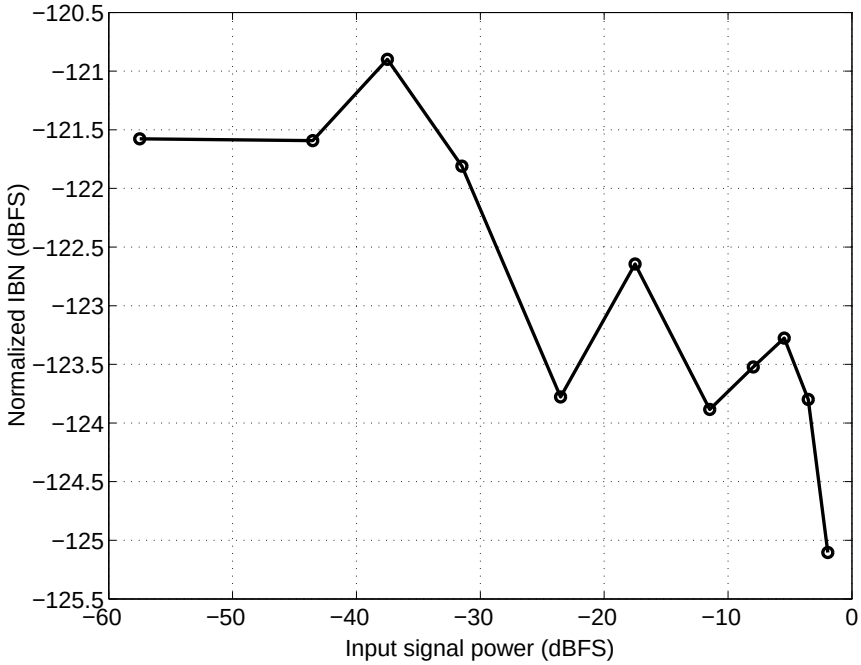


Figure 2.24: Variation of IBN with varying input signal power.

Fig. 2.24 shows the variation of in-band noise power (IBN) observed for varying input signal amplitudes. The IBN is seen to vary for very low amplitudes due to the nonlinear behavior of the quantizer at such low inputs. Otherwise, the noise power can be observed to be relatively independent of the variation in the signal power. Therefore, all the analytical estimates of IBN in this section were obtained using eq. 2.55, even though they were derived for a zero input modulator. $S_{xx}(\omega)$ is obtained as $NTF(e^{j\omega})\sigma_q^2/(2\pi)$, where σ_q^2 denotes the quantization noise power given as $\Delta^2/12$. $S_{zz}(\omega)$ is obtained through circular convolution of $S_{xx}(\omega)$.

Fig. 2.25 shows the analytically determined and simulated variation of IBN as a function of the OBG. As expected from the discussion in the previous section, as the OBG is increased, the total IBN first decreases and then increases, resulting in an op-

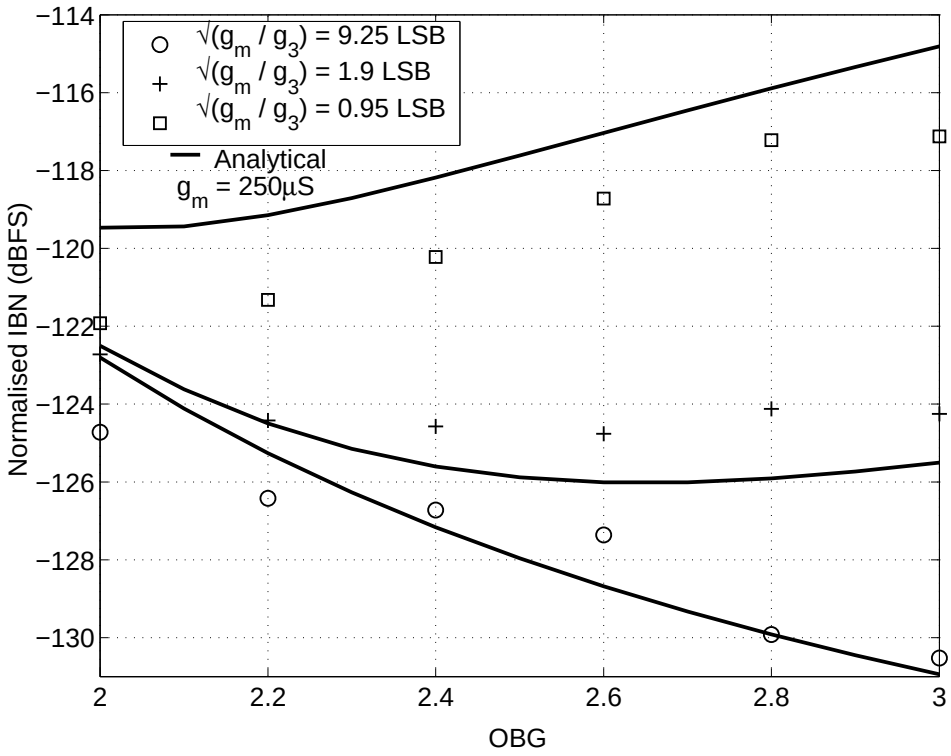


Figure 2.25: Variation of IBN with OBG for different levels of transconductor nonlinearity. The solid lines are the analytically computed results, while the points denote results from macromodel simulations.

timum OBG for a given nonlinearity. When the IBN is limited by noise due to nonlinearity ($\sqrt{g_m/g_3}$ = 0.95 LSB), the predicted IBN is found to deviate from the simulated noise by around 2 dB. This is a consequence of the deviation of the distribution of $x[n]$ from being truly Gaussian, as depicted in Fig. 2.15. In particular, the over-prediction of IBN can be attributed to the fact that $x[n]$ does not assume values as high as a truly Gaussian sequence would normally do. Further, as discussed before, this deviation in IBN is found to increase as the OBG is reduced.

Fig. 2.26 shows the variation of IBN with nonlinearity, as a function of the number of quantizer levels. The x-axis is $\sqrt{g_m/g_3}$, normalized to the LSB of the 4-bit quantizer. For this simulation, a maximally flat NTF was chosen to have optimally placed complex zeros, with OBG = 3. Consistent with our observations and predictions, we see that in-

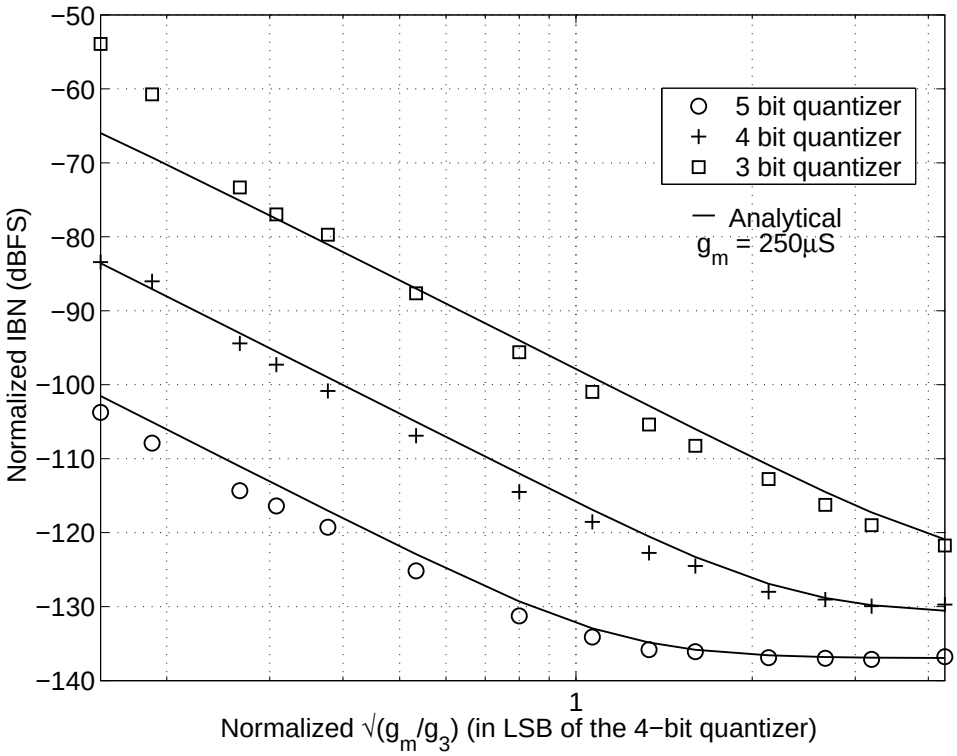


Figure 2.26: IBN as a function of nonlinearity for different quantizer resolutions.

creasing the quantizer resolution by one bit results in an 18 dB reduction in noise, when the performance of the modulator is nonlinearity dominated. When the nonlinearity is negligible, the performance is quantization noise limited and thus, the IBN reduces by 6 dB for every extra quantizer bit. While the analytically determined IBN is higher than the simulated values for reasons explained above, the trend is found to reverse at very high nonlinearities. This can be attributed to the failure of the assumption of weak nonlinearity, used while solving the cubic equation involving v_d. The condition $g_m v_d \gg g_3 v_d^3$ gets violated as the nonlinearity worsens, pushing the transconductor into strong nonlinear regions. This manifests as an unexpected increase in the IBN observed in simulations. For a given nonlinearity, such a behavior is understandably more evident with a 3-bit quantizer.

Fig. 2.27 shows the variation of IBN with nonlinearity for two values of g_m of the

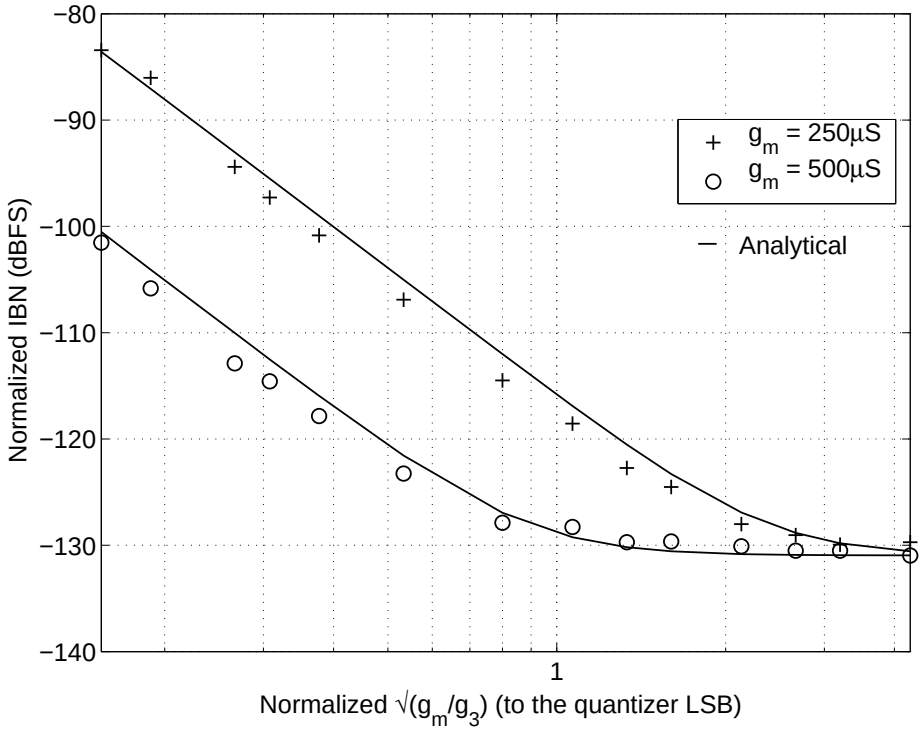


Figure 2.27: Variation of IBN with nonlinearity for two values of g_m.

transconductor used in the active-RC integrator. Again, for small values of $\sqrt{g_m/g_3}$, we see that halving g_m results in an increase in the IBN by 18 dB. This makes sense due to the following. From the discussion on modeling of integrator nonlinearity, if $g_m R \gg 2$ the value of $\beta \propto g_m^{-3}$. Using this in eq. 2.55, it can be seen that the IBN varies inversely as g_m^6, increasing by 18 dB when g_m is halved. As expected, the assumption of weak nonlinearity ($g_m v_d \gg g_3 v_d^3$) begins to fail at relatively smaller nonlinearities for the case where g_m = 250 μS, leading to an error in the IBN estimated.

For comparison purposes, the effect of nonlinearity in the single-stage integrator was observed for a CTDSM with RZ DAC. Fig. 2.28 compares the PSD of the modulator output with RZ and NRZ DAC for an input signal level of A_{in}=-37.5 dBFS. As predicted, the IBN is worse for the RZ case by 12 dB. Fig. 2.29 shows the dependency of the IBN in the RZ modulator with input signal power, for different levels of nonlinearity. The increase in IBN was predicted using eq. 2.71, by computing the power

spectral densities of the noise sequences using MATLAB.

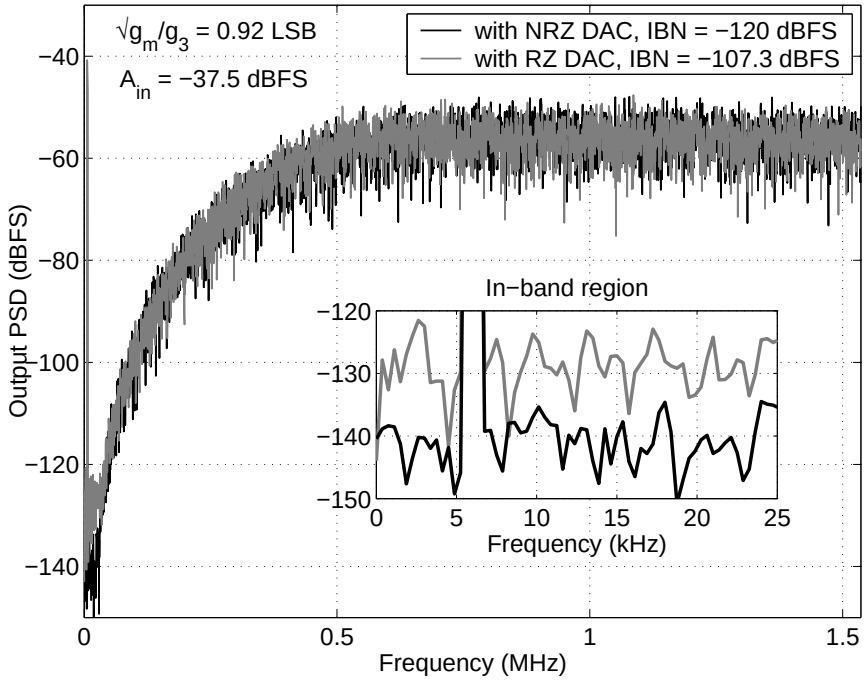


Figure 2.28: Comparison of output PSD of NRZ and RZ modulator with a nonlinear single-stage active-RC integrator (g_m=500 μS).

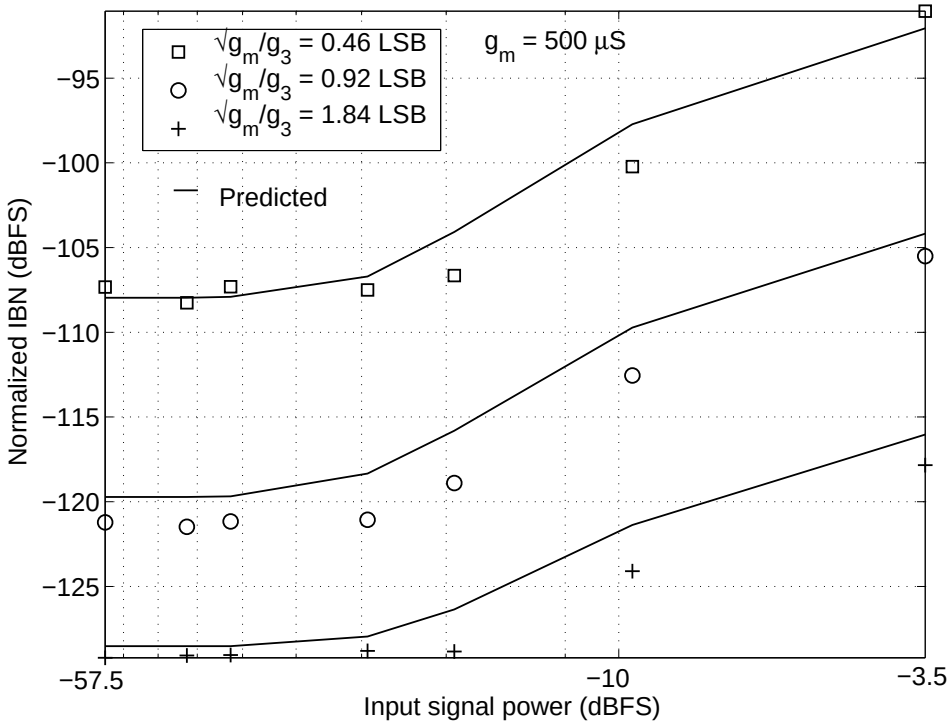


Figure 2.29: Variation of IBN with an RZ DAC, for varying input signal power and nonlinearity in the single-stage transconductor based active-RC integrator.

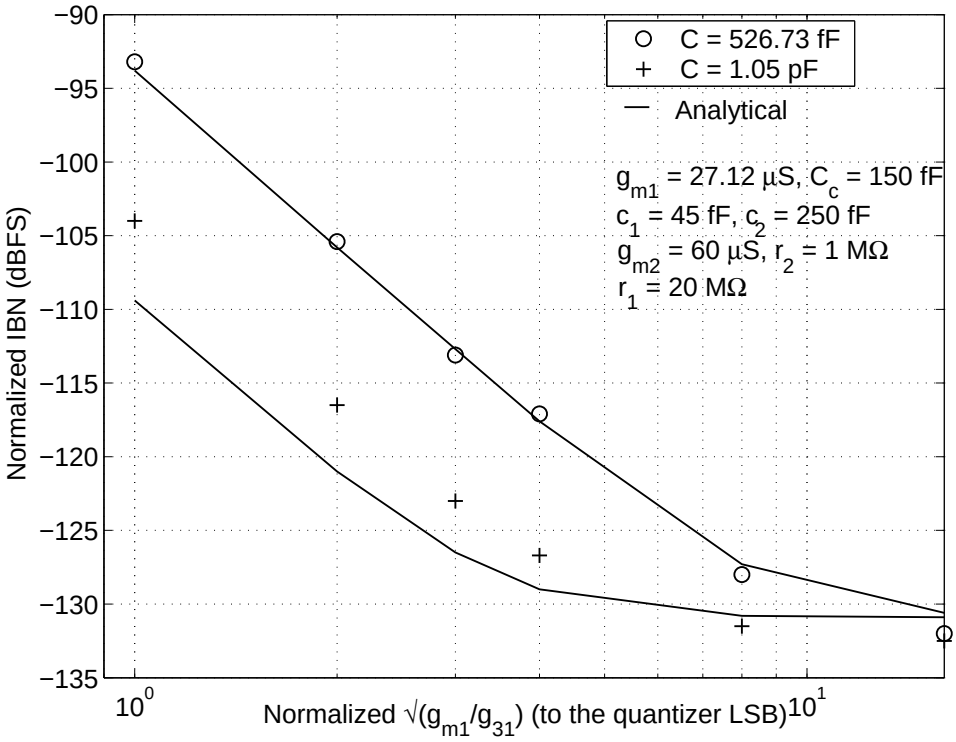


Figure 2.30: Variation of IBN with a two-stage Miller compensated opamp due to non-linearity in the input transconductor.

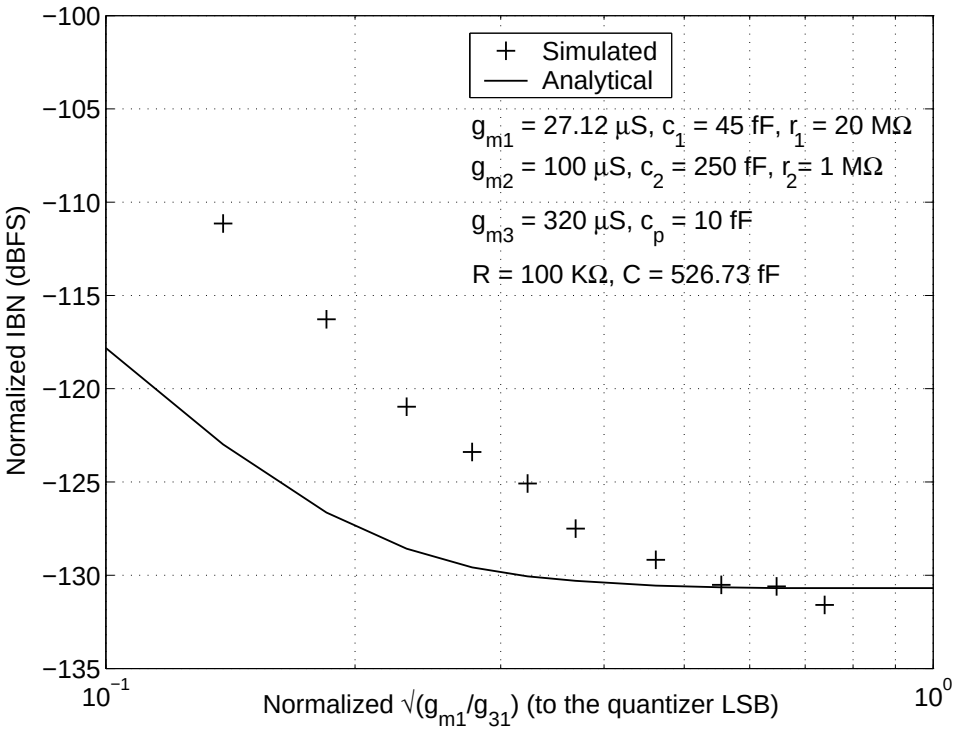


Figure 2.31: Variation of IBN with a feedforward compensated opamp.

Moving on to the results obtained with two-stage active-RC integrators, Fig. 2.30 shows the IBN observed (with an NRZ DAC) when a two-stage Miller opamp based integrator (see Fig. 2.5) is used. The model parameters of the integrator (g_{m1}, g_{m2}, C, C_c) are also indicated. Simulation results for two values of the integrating capacitor have been provided. The IBN determined using analytical expressions (eq. 2.55) closely matches those from macromodel simulation. Further, as expected from the discussion before, it is seen that increasing the integrating capacitance results in reduced IBN.

Finally, the performance of the modulator was observed with an integrator built with a feedforward compensated opamp having the same input transconductor as the Miller opamp. Fig. 2.31 shows the simulated and expected IBN for different nonlinearities of the input transconductor, along with the parameters of the integrator. It can be clearly seen from Fig. 2.30 that for the same nonlinearity, feedforward opamp based integrator gives better performance, thanks to the smaller value of c_p. However, note that there is significant error in the estimation of IBN for the feedforward opamp case. As briefly pointed out before, the cause for this under-estimation of the noise is the fact that, while deriving the nonlinear model for the integrator, the initial transients have been neglected in the response at the virtual ground node.

Illustrating the approximation, Fig. 2.32(a) shows the approximate and exact step response at v_d for the feedforward opamp based integrator. It can be seen that though the initial transients quickly die, they span a wider range of values, in comparison to the response without any transients. In a nonlinear integrator, the input transconductor can potentially be pushed into nonlinear regions during this period. The estimation of IBN neglecting the initial transients can therefore be in error, if the nonlinear currents

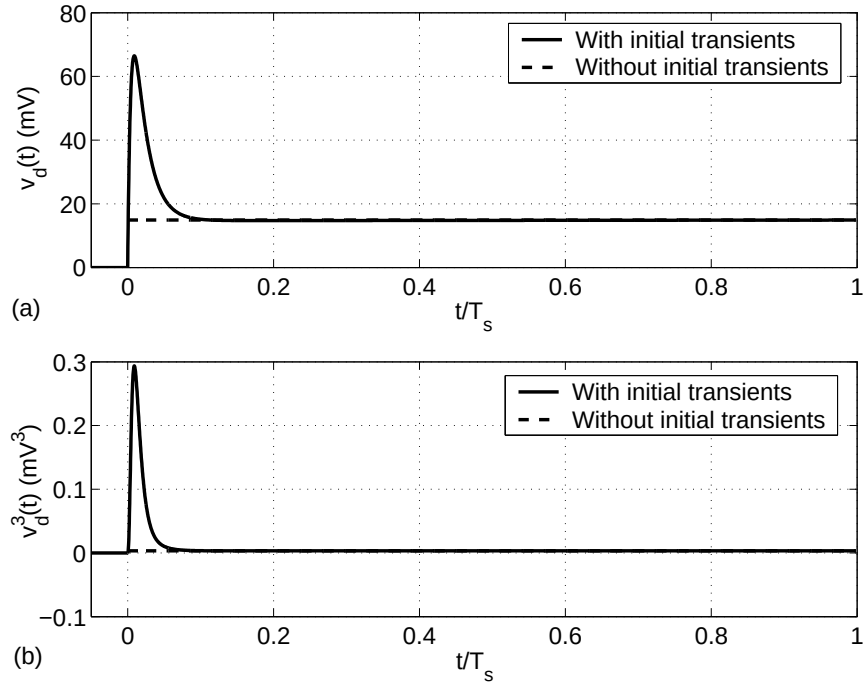


Figure 2.32: Response at the virtual ground node of the active-RC integrator with feed-forward compensated opamp, to a step input. (a) $v_d(t)$ (b) $v_d^3(t)$

injected by the input transconductor (i_{inj}) are considerably higher during this transient period (as seen from $v_d^3(t)$ shown in Fig. 2.32(b)). For the same reason, a similar under-estimation can be seen in Fig. 2.30 for the Miller opamp case with C = 1.05 pF. In all these cases, because of a large value for k (obtained by increasing C or reducing C_c or c_p), one arrives at a situation where the initial transients[13] begin to constitute a significant part of the response at v_d. The aim of the following chapter is to address this issue, using the intuition developed in this chapter.

[13]The initial transients largely depend on g_{m3}, g_{m2}, c_1 and c_2 and hence, are not affected much due to a change in the factor k.

CHAPTER 3

MODELING THE EFFECT OF INITIAL TRANSIENTS

In the analysis of Chapter 2, the initial transients at the virtual ground of a two-stage opamp based active-RC integrator were neglected. This resulted in an under-estimation of the in-band noise due to integrator nonlinearity, especially for the case of the feed-forward opamp based active-RC integrator. This chapter is aimed at modeling the effect of initial transients on the performance of multi-bit CTDSMs.

In essence, with the input to the first integrator being a sampled and held version of the shaped quantization noise, the relevant step response at the virtual ground node (v_d) can be considered as having two components - a steady state part of the form $G_0u(t)$ and a transient part, $w(t)$. As shown in Fig. 3.1, we thus have

$$v_d(t) \approx G_0u(t) + w(t) \tag{3.1}$$

where $w(t)$ can be assumed to decay to zero beyond $t = t_1$, with $t_1 \ll T_s$ in a fast opamp.

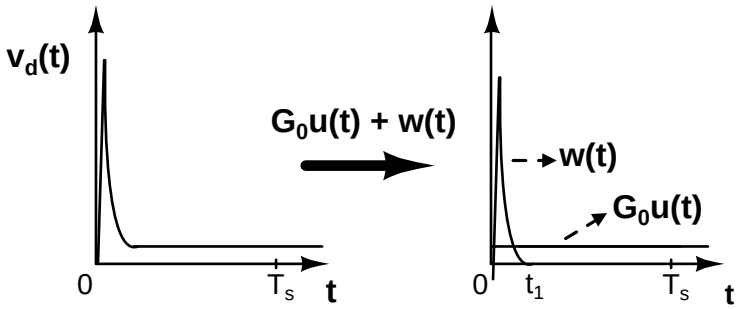


Figure 3.1: Step response at the virtual ground node of an active-RC integrator with two-stage opamp, considered as having two components - $G_0u(t)$ and $w(t)$.

It can be recalled that the gain factor G_0 was intuitively derived to be $1/(2 + kg_{m1}R)$, where k is appropriately obtained for Miller and feedforward compensated opamps. Proceeding along similar lines as in Chapter 2, the effect of initial transients can be modeled for multi-bit modulators in the following manner.

3.1 Modeling of integrator nonlinearity including the effect of initial transients

Consider the case of a feedforward opamp based active-RC integrator used in a multi-bit CTDSM with an NRZ DAC, as shown in Fig. 3.2. The effect of weak nonlinearity in the input transconductor can be analyzed using Bussgang's method of current injection.

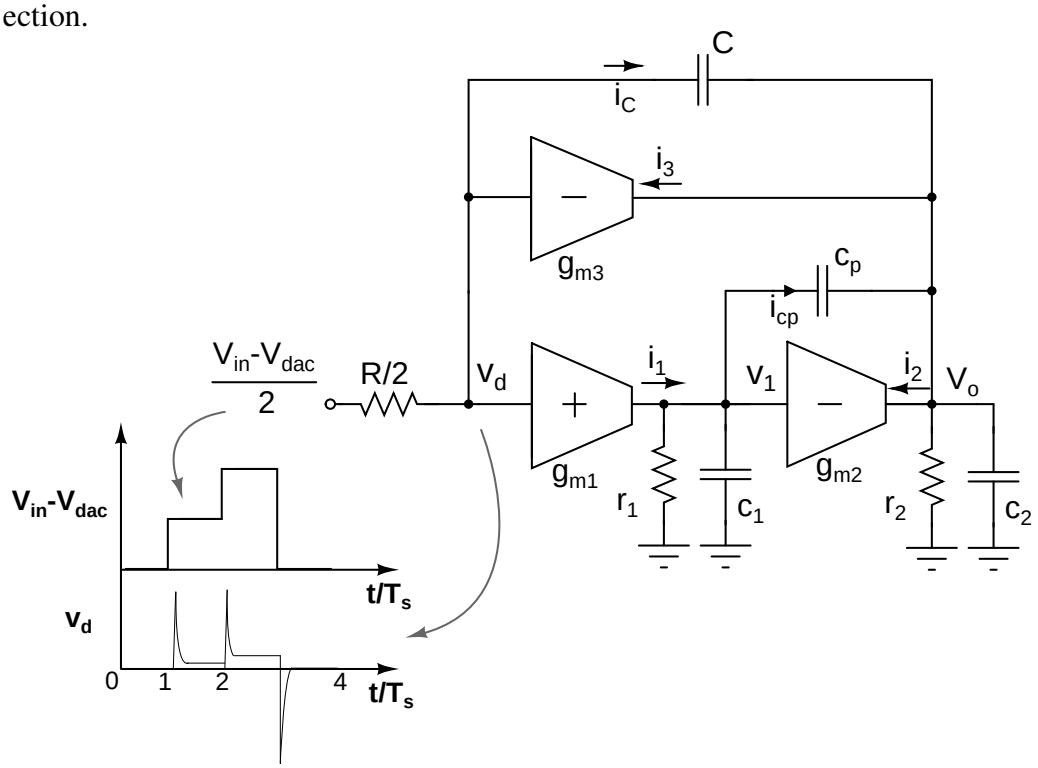


Figure 3.2: Circuit of an active-RC integrator with a two-stage feedforward compensated opamp used in a multi-bit CTDSM with an NRZ DAC.

The first task involves the determination of the linear response at the virtual ground node, v_d. As illustrated in Fig. 3.2, the input to the integrator is a series of step-like

signals with an amplitude $x[n] = V_{in}[n] - V_{out}[n]$ in each clock cycle. Denoting the first difference of the shaped quantization noise sequence as $x'[n] = x[n] - x[n-1]$, the linear response $v_{d,lin}(t)$ is thus given by

$$v_{d,lin}(t) \approx G_0 x[n] + x'[n] w(t - nT_s), \quad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s} \tag{3.2}$$

The next task is to determine the injected response, $v_{d,inj}(t)$. With a weakly nonlinear input transconductor ($i_1 = g_{m1} v_d - g_{31} v_d^3$), this involves the injection of the current $i_{inj} = g_{31} v_{d,lin}^3$ as shown in Fig. 3.3.

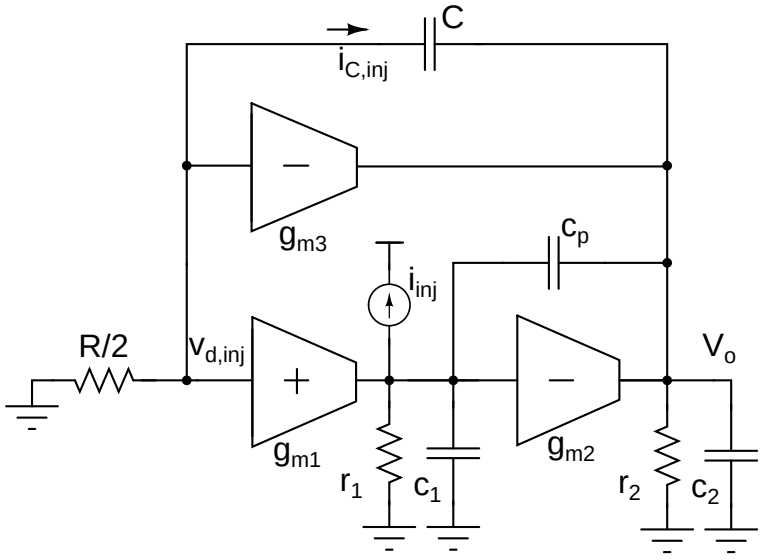


Figure 3.3: Injection of nonlinear current (i_{inj}) due to the input transconductor in an active-RC integrator with two-stage feedforward compensated opamp.

From eq. 3.2, the injected current in a given clock cycle ($\mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s}$) can be expressed as

$$i_{inj}(t) \approx g_{31}\Big(G_0 x[n] + x'[n] w(t - nT_s)\Big)^3 \tag{3.3}$$

$$\begin{aligned} = \; & g_{31} G_0^3 x^3[n] + 3 g_{31} G_0^2 x^2[n] x'[n] w(t - nT_s) \\ & + 3 g_{31} G_0 x[n] x'^{\,2}[n] w^2(t - nT_s) \\ & + g_{31} x'^{\,3}[n] w^3(t - nT_s), \qquad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s} \end{aligned} \tag{3.4}$$

If initial transients are neglected, the first term in the above expression leads us to the familiar nonlinear function, $f(x) = x - \beta x^3$ (as derived in Chapter 2). This is made possible because of a "memoryless" relation between $x[n]$ and the injected current for each clock cycle. However, the presence of the initial transients ($w(t)$) in the above expression does not permit such a relation, thus complicating the analysis. We overcome this difficulty as discussed below.

It can be seen that the injected current will appear at the output of the integrator (and the modulator output) after passing through a linear network[1]. In essence, there are no more nonlinear operations in its path that can fold back (or demodulate) the out-of-band noise into the in-band region. This implies that the knowledge of the "low frequency" content of $i_{inj}(t)$ is sufficient to determine the in-band noise due to integrator nonlinearity.

The spectrum of the injected current can be found by considering it as having four random components that are filtered versions of the shaped quantization noise ($x[n]$) and its first difference ($x'[n]$). For instance, the last term in eq. 3.4 can be perceived as the output of a filter with impulse response given by $w^3(t)$ and whose input is a train of impulses scaled by the samples, $g_{31}x'^{\,3}[n]$. Thus, denoting the PSD of the injected current as $S_{ii}(f)$, it can be found at low frequencies ($f \ll f_s$) as (Papoulis and Pillai (2002))

$$\begin{aligned} S_{ii}(f) \;\approx\; & T_s(g_{31}G_0^3)^2 S_{x^3x^3}(f) + \frac{1}{T_s}\left|3g_{31}G_0^2\int_0^{t_1} w(t)\ \mathrm{dt}\right|^2 S_{x^2x'}(f) \\ & + \frac{1}{T_s}\left|3g_{31}G_0\int_0^{t_1} w^2(t)\ \mathrm{dt}\right|^2 S_{xx'^2}(f) + \frac{1}{T_s}\left|g_{31}\int_0^{t_1} w^3(t)\ \mathrm{dt}\right|^2 S_{x'^3}(f) \\ & + \text{cross correlation terms}\,, \quad (\text{for f} \ll \text{f}_\text{s}) \end{aligned} \tag{3.5}$$

[1]With the usual assumption that the quantizer is linear.

where components like $S_{x'^3}(f)$ are periodic replicas of $S_{x'^3}(\omega)$ repeated every multiples of f_s. Assuming a sampled and held injected current (say obtained from the samples, $i_{inj,eq}[n]$), it will have a low frequency PSD given by $T_s S_{i_{eq}i_{eq}}(f)$. Comparing this with eq. 3.5, the low frequency content will be same as that of $S_{ii}(f)$, if

$$S_{i_{eq}i_{eq}}(f) = \frac{S_{ii}(f)}{T_s}, \quad (\text{for f} \ll \text{f}_\text{s}) \tag{3.6}$$

$$\begin{aligned} = & (g_{31}G_0^3)^2 S_{x^3x^3}(f) + \frac{1}{T_s^2}\left|3g_{31}G_0^2\int_0^{t_1} w(t)\ \mathrm{dt}\right|^2 S_{x^2x'}(f) \\ & + \frac{1}{T_s^2}\left|3g_{31}G_0\int_0^{t_1} w^2(t)\ \mathrm{dt}\right|^2 S_{xx'^2}(f) + \frac{1}{T_s^2}\left|g_{31}\int_0^{t_1} w^3(t)\ \mathrm{dt}\right|^2 S_{x'^3}(f) \\ & + \frac{1}{T_s^2}[\text{cross correlation terms}] \end{aligned} \tag{3.7}$$

Thus by defining,

$$\alpha_2 = \frac{1}{T_s}3G_0^2\int_0^{t_1} w(t)\ \mathrm{dt} \tag{3.8}$$

$$\alpha_3 = \frac{1}{T_s}3G_0\int_0^{t_1} w^2(t)\ \mathrm{dt} \tag{3.9}$$

$$\alpha_4 = \frac{1}{T_s}\int_0^{t_1} w^3(t)\ \mathrm{dt} \tag{3.10}$$

the injected current can be modeled as having four sampled and held discrete-time sequences given by[2]

$$i_{inj,eq}[n] \approx g_{31}\Big(G_0^3x^3[n] + \alpha_2x^2[n]x'[n] + \alpha_3x[n]x'^2[n] + \alpha_4x'^3[n]\Big) \tag{3.11}$$

Physically, the low frequency equivalence can be attributed to the same charge being injected by both $i_{inj}(t)$ and the above currents, in each clock cycle. The expressions for α_i (i=2,3,4) can therefore be seen as the time-average of the charges transferred by the respective transient terms present in eq. 3.4, in a given clock cycle.

[2]The reader can refer to Appendix D for a detailed mathematical proof on the above derivation.

Given a series of step currents generated from the above sequence, the results derived in the previous chapter (Section 2.1.2) can be used to determine the injected response at v_d. Since we are only interested in the low frequency content, it is fair to assume that the step-like injected currents result in a series of step signals[3] at v_d. Thus, the equivalent injected response ($v_{d,inj,eq}$) in a given clock cycle and the resulting capacitive current ($i_{C,inj,eq}$) can be found as

$$v_{d,inj,eq} = \frac{kR}{(2+kg_{m1}R)} i_{inj,eq} \quad \text{and so,} \tag{3.12}$$

$$i_{C,inj,eq} = \frac{-2v_{d,inj,eq}}{R} = \frac{-2k}{(2+kg_{m1}R)} i_{inj,eq} \tag{3.13}$$

In a linear integrator, the capacitor current in each clock cycle is related to the integrator input as

$$i_C = \frac{kg_{m1}}{(2+kg_{m1}R)}(V_{in} - V_{out}) \tag{3.14}$$

Using this relation, $i_{C,inj,eq}$ can be generated through an additional input sequence to the linear integrator given by

$$z[n] = \frac{-2i_{inj,eq}[n]}{g_{m1}} \tag{3.15}$$

thus resulting in the model for the nonlinear integrator[4] as shown in Fig. 3.4.

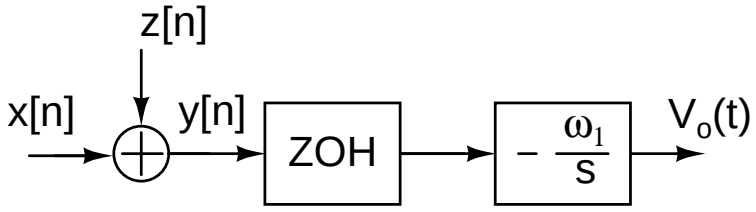


Figure 3.4: Model of the nonlinear active-RC integrator with $z[n] = \sum_{i=1}^{4} z_i[n]$, that includes the effect of initial transients in the in-band region.

[3] neglecting the initial transients

[4] The model can be used to analyze the effect of integrator nonlinearity in the in-band region.

Assuming the random sequence ($z[n]$) to consist of four sequences, $z_i[n]$ (i=1,2,3,4), $y[n]$ is given by

$$y[n] = x[n] + \sum_{i=1}^{4} z_i[n] \tag{3.16}$$

$$= x[n] - \beta_1 x^3[n] - \beta_2 x^2[n]x'[n] - \beta_3 x[n]x'^2[n] - \beta_4 x'^3[n] \tag{3.17}$$

where,

$$\beta_1 = \frac{2g_{31}G_0^3}{g_{m1}} = \frac{2g_{31}}{g_{m1}(2+kg_{m1}R)^3}$$

$$\beta_i = \frac{2g_{31}\alpha_i}{g_{m1}} = \beta_1\frac{\alpha_i}{G_0^3}, \quad (\text{for i} = 2,3,4)$$

It can be seen that the nonlinear coefficient, β_1 is same as that derived in Chapter 2 (Section 2.1.2) with the other coefficients being easily computed from the knowledge of the time-averages (α_i).

Though the nonlinear model for the integrator was derived by considering the nonlinearity of the input transconductor in a feedforward opamp based integrator, a similar analysis can be done for a nonlinear feedforward transconductor and also for a nonlinear input transconductor in a Miller opamp based integrator. Using the respective β_1 and the time-averages, the in-band noise due to nonlinearity can be determined in the same manner as in eq. 3.20.

3.2 Modeling the effect of initial transients in a multi-bit CTDSM

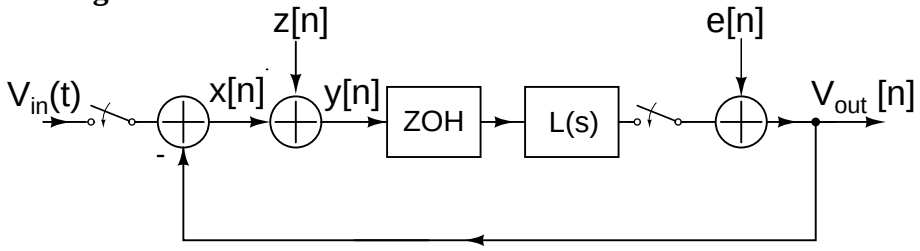


Figure 3.5: Model of the CTDSM with a nonlinear integrator where $z[n] = \sum_{i=1}^{4} z_i[n]$ includes the effect of initial transients. The model can be used to predict the in-band noise due integrator nonlinearity.

The in-band noise due to integrator nonlinearity in a multi-bit CTDSM can be determined from the model for the CTDSM as shown in Fig. 3.5, where $z[n]$ is now considered to have four sequences. The autocorrelation sequence of $y[n]$ is therefore

$$R_{yy}[n] = R_{xx}[n] + \sum_{j=1}^{4}\sum_{m=1}^{4} R_{z_j z_m}[n] \tag{3.18}$$

where $R_{z_j z_m}[n]$ denotes the cross-correlation computed for the sequences $z_j[n]$ and $z_m[n]$. Equivalently, the corresponding PSD can be expressed as

$$S_{yy}(\omega) = S_{xx}(\omega) + \sum_{j=1}^{4}\sum_{m=1}^{4} S_{z_j z_m}(\omega) \tag{3.19}$$

$$= S_{xx}(\omega) + \sum_{j=1}^{4} S_{z_j z_j}(\omega) + \sum_{j=1}^{4}\sum_{m=1,m>j}^{4} 2Re\left(S_{z_j z_m}(\omega)\right) \tag{3.20}$$

where $S_{z_j z_m}(\omega)$ denotes the cross power spectral density and $Re\left(S_{z_j z_m}(\omega)\right)$ denotes the respective real part. When compared to the expression in eq. 2.55, the inclusion of initial transients is seen to have introduced additional spectral components because of the new sequences $z_2[n]$, $z_3[n]$ and $z_4[n]$. Given the samples of the shaped quantization noise[5] $x[n]$, the (cross) spectral densities can be determined using tools like MATLAB to quantify the expected in-band noise of the multi-bit CTDSM.

While such numerical computations can accurately determine the performance of the modulator, some useful analytical insights can be obtained by approximating the in-band noise of the modulator as

$$S_{yy}(\omega) \approx S_{xx}(\omega) + S_{z_1 z_1}(\omega) + S_{z_4 z_4}(\omega) \tag{3.21}$$

[5]These samples can be obtained by passing uniform/Gaussian quantization noise through the NTF.

Assuming the sequences $x[n]$ and $x'[n]$ to be Gaussian processes,

$$S_{z_1z_1}(\omega) = \mathcal{F}\Big[6\beta_1^2 R_{xx}^3[n]\Big] \tag{3.22}$$

$$S_{z_4z_4}(\omega) = \mathcal{F}\Big[6\beta_4^2 R_{x'x'}^3[n]\Big] \tag{3.23}$$

Defining $S_{x'x'}(\omega) = |1 - e^{-j\omega}|^2 S_{xx}(\omega)$, the above spectral components can be determined through circular convolution of $S_{xx}(\omega)$ and $S_{x'x'}(\omega)$. Given the NTF and the details of nonlinearity (β_1 and β_4 as computed from g_{31}, g_{m1}, G_0 and $w(t)$), the in-band noise due to integrator nonlinearity can thus be analytically determined using eq. 3.21. Understandably, these estimates will be accurate for cases where the initial transients are negligible ($w(t) \approx 0$) or when the operating speed of the modulator (f_s) is small. In such cases, the in-band noise will be purely determined by $S_{xx}(\omega)$ and $S_{z_1z_1}(\omega)$ as derived in Chapter 2. Similarly, in cases where $G_0 \approx 0$ (like the feed-forward opamp based integrator), the initial transients become significant. This gets translated to $\beta_1 = \beta_2 = \beta_3 \approx 0$, thus resulting in $S_{yy}(\omega) \approx S_{z_4z_4}(\omega)$. In these cases, due to the enhanced high frequency components of $x'[n]$ (when compared to that of $x[n]$), one can expect larger demodulation of the out-of-band noise into the signal band. This results in greater performance degradation due to integrator nonlinearity.

3.3 Simulation results

The accuracy of the model derived by taking into account the initial transients was verified using the same third order CTDSM (as in Chapter 2) having a 4-bit quantizer. The validity of approximating the individual transient components by their time-averages was first confirmed. Fig. 3.6 compares the spectrum of $v_{d,lin}^3(t)$ obtained through simulation and the spectrum of $v_{d,lin,eq}^3(t)$ (=$i_{inj,eq}(t)/g_{31}$ from eq. 3.11) de-

termined using the time-averages. The inset shows good matching between the two spectra in the in-band region[6], thus validating the approximation.

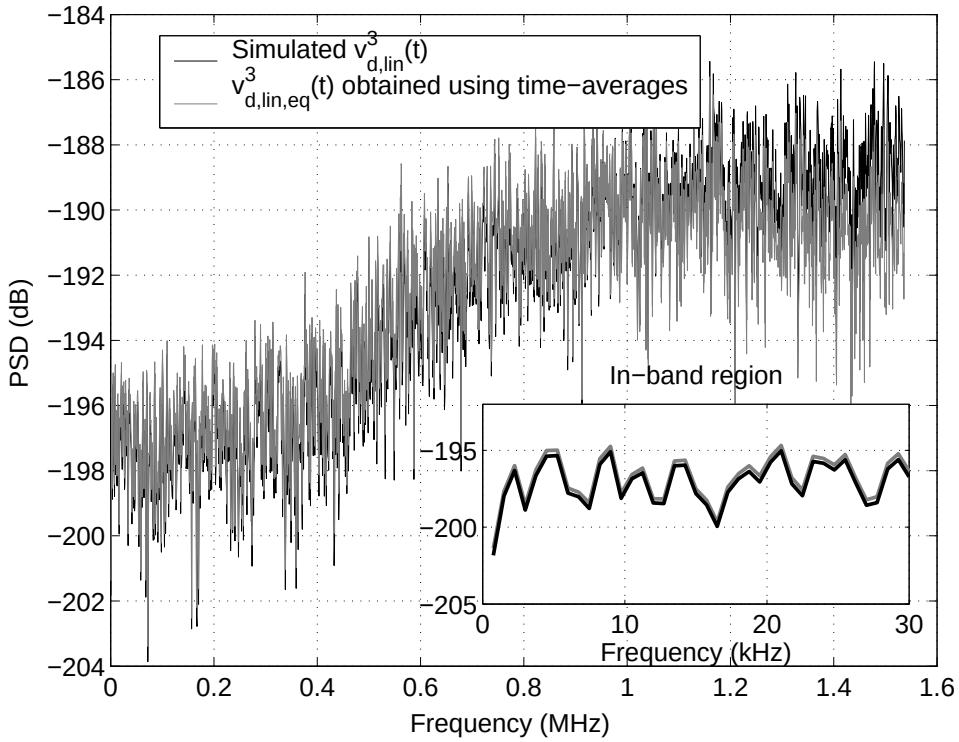


Figure 3.6: Comparison between the spectrum of the exact and approximate version of $v_{d,lin}(t)^3$ obtained with a feedforward opamp based integrator. g_{m1} = 27.12 μS, R = 100 kΩ, C = 527 fF, c_1 =45 fF, c_p =10 fF, g_{m2} = 100 μS, c_2 =250 fF, g_{m3} = 320 μS, r_2 = 1 MΩ. The inset shows the matching between the two spectra in the in-band region.

Having confirmed this, the in-band noise powers expected for the feedforward opamp based integrator (simulated in Chapter 2) were recomputed and compared with that obtained through simulations. The noise powers were calculated through eq. 3.20, where each one of the spectral components was computed in MATLAB, using shaped quantization noise samples $x[n]$ obtained from Gaussianly[7] distributed $e[n]$. The relative values of the nonlinear coefficients (β_i/β_1) used for calculations are as given in Ta-

[6]The sampled and held natured of $v^3_{d,lin,eq}(t)$ is the reason for the deviation of the spectrum at higher frequencies due to the $sinc(fT_s)$ function.

[7]Since we are working with time-samples, a uniform distribution can also be assumed to obtain $x[n]$. However, for uniformity with the results previously derived, a Gaussian distribution is used. Note that this results in an over-estimation of the noise by about 2 dB.

Parameter changed	β_2/β_1	β_3/β_1	β_4/β_1
–	0.28	0.65	0.59
g_{m3} = 640 μS	0.22	0.23	0.09
g_{m3} = 480 μS	0.25	0.37	0.22
c_1 = 90 fF	0.5	1.11	1.02
c_1 = 25 fF	0.17	0.42	0.36
c_2 = 400 fF	0.34	0.95	1
c_2 = 100 fF	0.21	0.4	0.3
g_{m2} = 200 μS	0.14	0.34	0.29
g_{m2} = 150 μS	0.19	0.45	0.4
C = 1.05 pF	0.5	2.05	3.32
C = 790 fF	0.39	1.26	1.6
R = 200 kΩ	0.28	0.68	0.64
R = 150 kΩ	0.28	0.67	0.62

Table 3.1: Relative nonlinear coefficients in feedforward opamp based integrator with g_{m1} = 27.12 μS, R = 100 kΩ, C = 526.73 fF, r_1 = 20 MΩ, c_p =10 fF, g_{m2} =100 μS, g_{m3} = 320 μS, c_1 =45 fF, r_2 = 1 MΩ

ble 3.1. Fig 3.7 compares the simulated IBN observed for the case of $g_{m3} = 320\,\mu$S with the IBN predicted with and without including the initial transients ($w(t)$). It is apparent that the IBN is more accurately predicted after including the effect of initial transients. Alternatively, the in-band noise can also be analytically estimated (using eq. 3.21) in terms of the components $S_{z_1z_1}(\omega)$ and $S_{z_4z_4}(\omega)$ (computed through circular convolution of $S_{xx}(\omega)$ and $S_{x'x'}(\omega)$). For illustrative purposes, the IBN estimated through this analytical method is also shown. It can be seen that the estimation is better than the case where $w(t)$ is completely neglected, though with a small inaccuracy with the simulation results. As seen from Table 3.1, this is because of the appreciable values of β_2 and β_3 arising due to the comparable significance of $w(t)$ and $G_0u(t)$ for the given integrator.

To ensure that the modeling is not limited to a specific combination of opamp parameters, exhaustive simulations were carried out for different combinations. Since

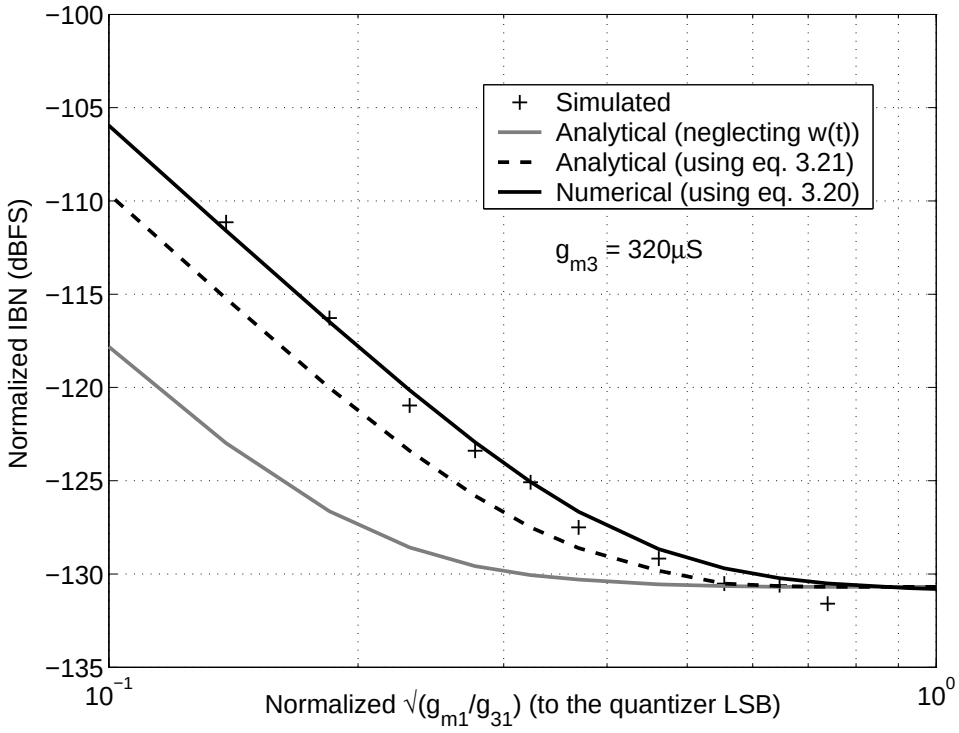


Figure 3.7: Comparison of the IBN estimated with and without including the effect of initial transients for the feedforward opamp based integrator.

$w(t)$ is not the only significant response in these integrators, the numerical computation method was used to predict the IBN in all these cases, for better accuracy. Fig. 3.8 shows the variation observed in the IBN for two values of g_{m3}. The performance improvement observed with increasing g_{m3} can be attributed to the reduction in the initial transients, as is evident from the trend shown in Table 3.1. Fig. 3.9 and 3.10 show the observed and predicted variation in the IBN for different values of c_1 and c_2, respectively. Intuitively, an increase in these capacitors can be expected to worsen the initial transients and therefore, the IBN is found to increase with increasing values of c_1 and c_2. The corresponding variation in the IBN with different integrating capacitors is shown in Fig. 3.11. It can be seen that the predicted results match closely with that obtained from simulations. Due to the presence of the initial transients, a marginal improvement is only seen when C is doubled, as against the expected reduction of 18 dB.

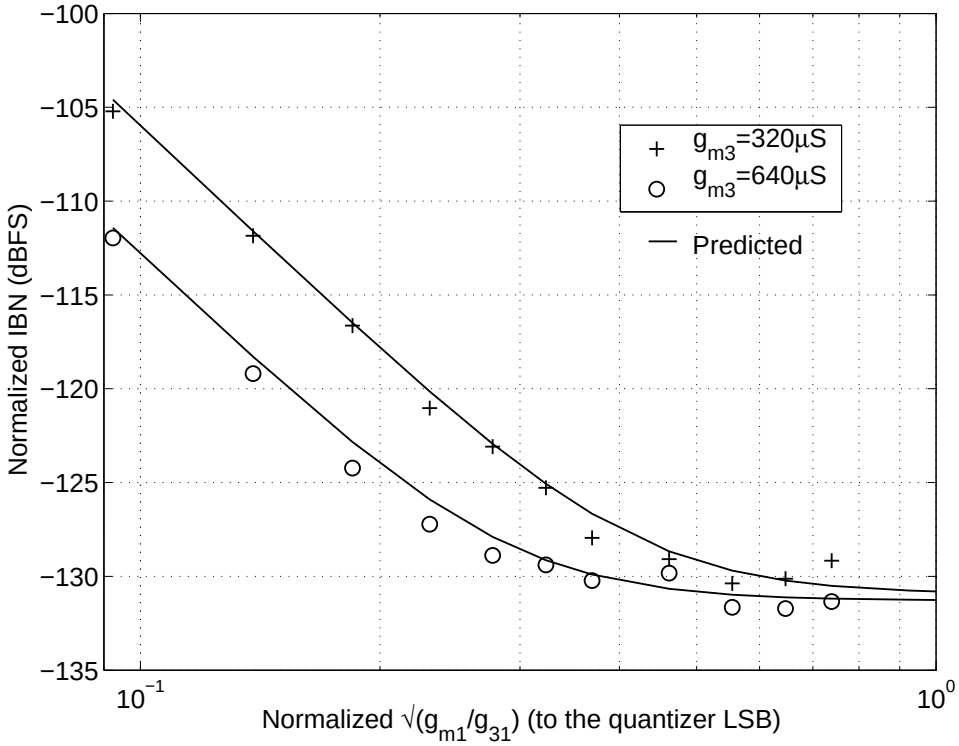


Figure 3.8: Variation in IBN with a feedforward opamp based integrator for two values of g_{m3}.

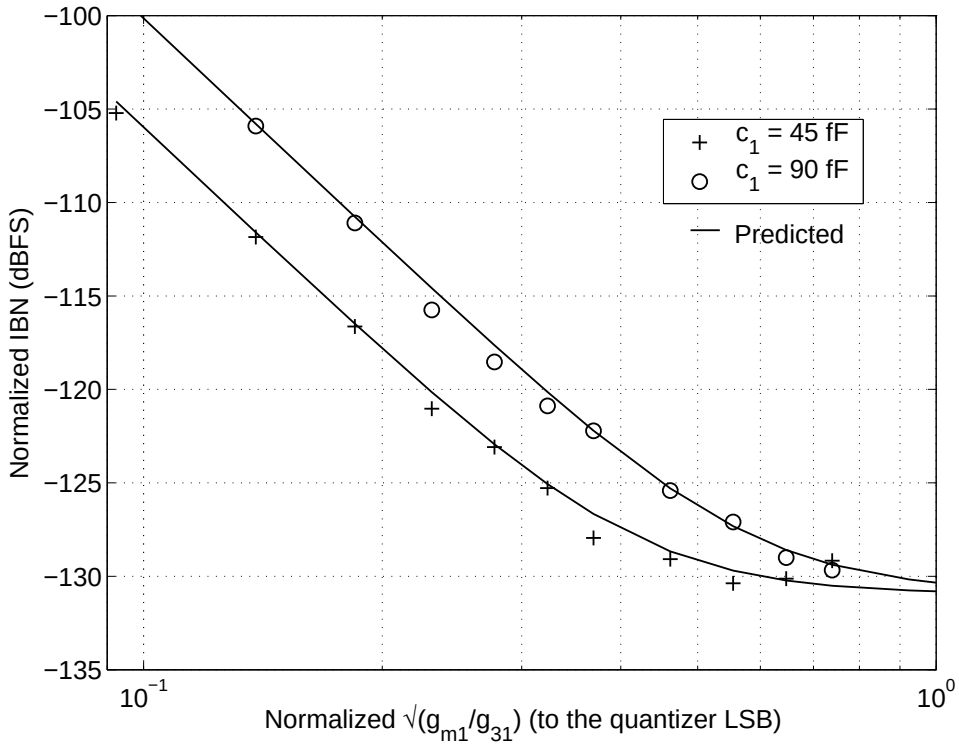


Figure 3.9: Variation in IBN with a feedforward opamp based integrator for two values of c_1.

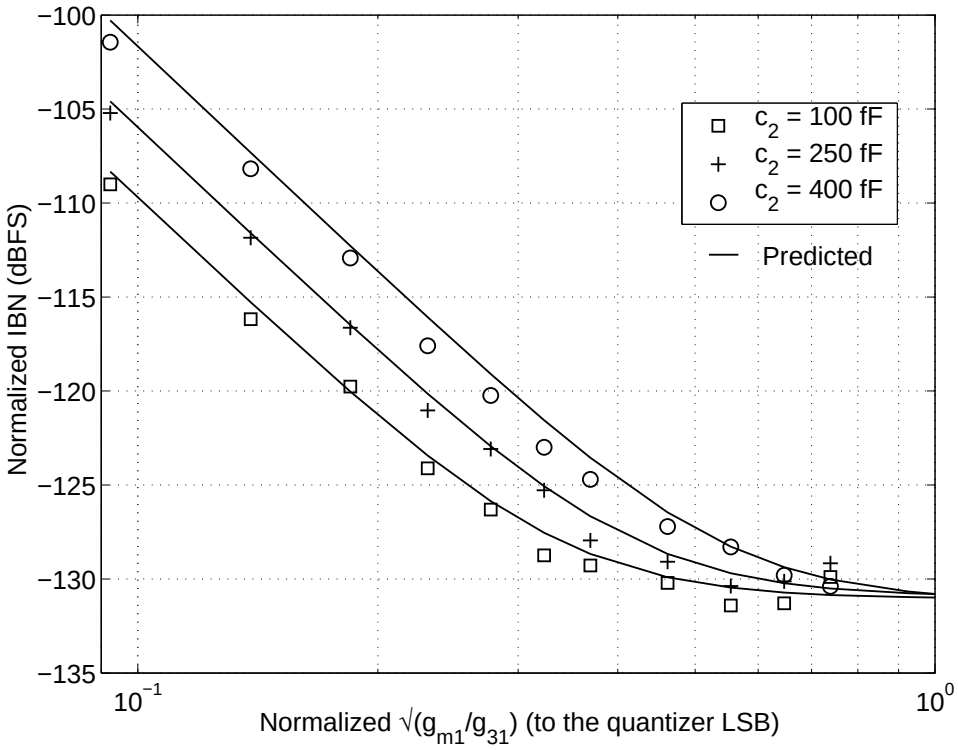


Figure 3.10: Variation in IBN with feedforward opamp based integrator for various c_2.

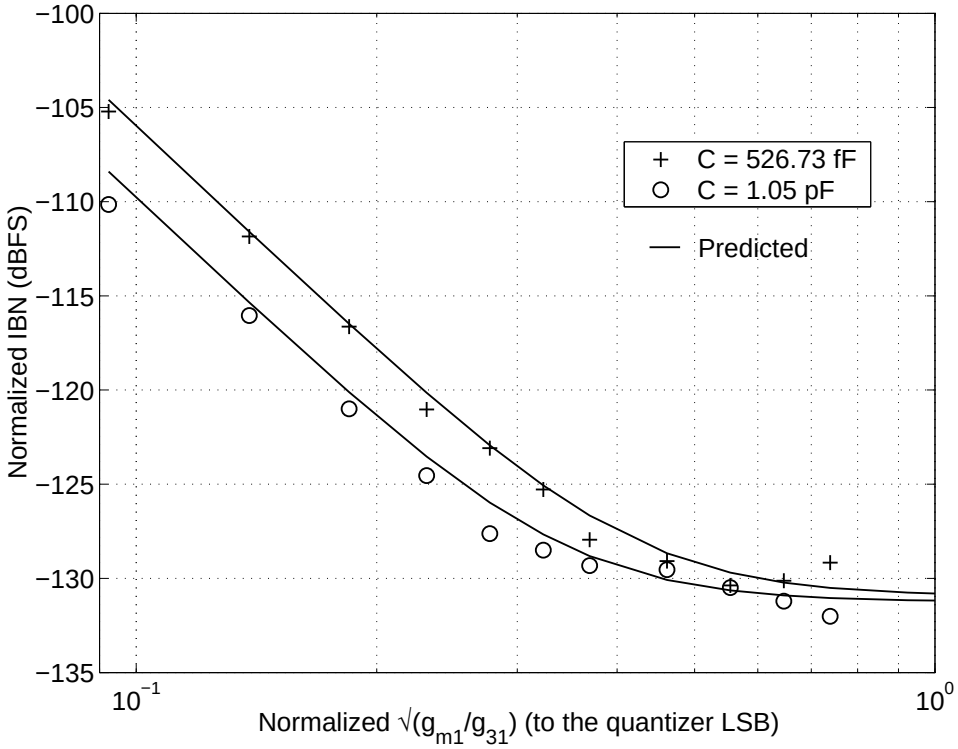


Figure 3.11: Variation in IBN with a feedforward opamp based integrator for two values of integrating capacitors.

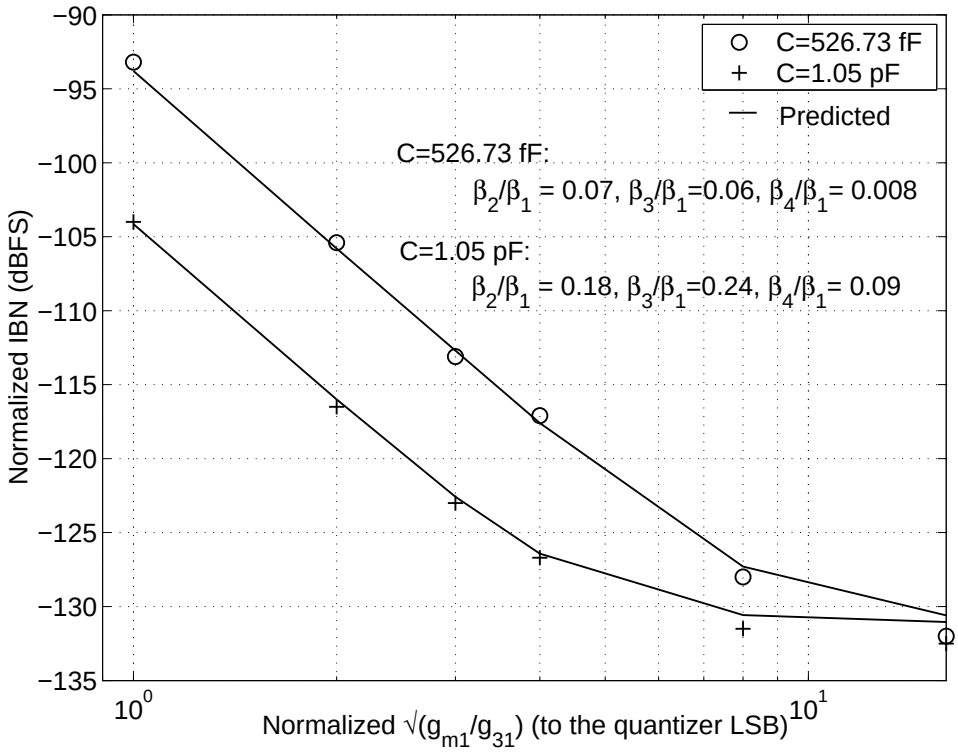


Figure 3.12: Variation in IBN with Miller opamp based integrator for two values of C.

The derived results were also verified for the Miller opamp based integrator used in Chapter 2. Fig. 3.12 shows the observed and predicted variation in the IBN for different integrating capacitors. The ratios β_i/β_1 have also been provided in the figure. As was shown in the previous chapter (Fig. 2.30), the effect of initial transients is negligible with C = 526.73 fF. However, the modeling of initial transients can be seen to predict the IBN more accurately for the case where C = 1.05 pF.

3.4 Extensions to the analysis

The discussions so far have analyzed the effect of integrator nonlinearity in multi-bit CTDSMs. Comparison of various integrator topologies showed that two-stage active-RC integrators, in particular those with feedforward compensation, are a good choice to adopt in high resolution CTDSMs. Though such comparisons were done by as-

suming a resistive feedback DAC, the reader can easily see that the corresponding nonlinear model with a current-steering DAC will in general be of the form $f(x) = x - \frac{g_3}{g_m(1+g_m kR)^3}$ for active-RC integrators. This signifies a 6 dB improvement in the overall performance with the use of a current-steering architecture. The final call on the choice of the DAC, however, also depends on others factors like noise, area and speed (as discussed in Appendix E).

The effect of initial transients can also be analyzed for modulators employing RZ DAC, similar to the NRZ case. However, the modulator performance with an RZ DAC is worse[8] than with an NRZ DAC by about 12 dB. As discussed in the previous chapter, such performance degradations further worsen with increasing input amplitudes. Moreover, as illustrated in Appendix F, the slew-rate requirements can be considerably higher in multi-bit CTDSMs when an RZ DAC is employed. In addition, even when slewing conditions are avoided, the increased magnitude of transitions in the integrator input[9] can enhance the initial transients at the virtual ground node. In such cases, the condition for weak nonlinearity assumed for the transconductors can be violated, which further increases the performance degradation. Given the increased severity of integrator nonlinearity with RZ DACs, the analysis with initial transients has not been shown in this chapter, for the sake of brevity.

So far, we have neglected the nonlinearity of the second stage transconductor in the analysis involving two-stage active-RC integrators. It turns out that the effect of such nonlinearities gets attenuated by the gain of the first stage. Though not shown in the chapter, simulations run with the two-stage integrators considered so far reveal that the

[8] even without considering the initial transients

[9] With an RZ DAC, the transition in each clock cycle is $2V_{out}[n]$. In contrast, the transitions span only 2 or 3 LSBs with an NRZ DAC.

effect of second stage nonlinearity is indeed negligible. For a quantitative understanding, the reader can refer to the discussion in Appendix G.

Having analyzed multi-bit modulators, the logical step to take further is to apply these concepts to model the effect of integrator nonlinearity in single-bit modulators.

CHAPTER 4

ANALYSIS OF INTEGRATOR NONLINEARITY IN CTDSMs : THE SINGLE-BIT CASE

In the previous chapters, integrator nonlinearity in multi-bit CTDSMs was shown to result in increased in-band noise. The performance degradation was found to worsen by 18 dB for every one bit reduction in the resolution of the quantizer. Therefore, a severe increase in the in-band noise floor can be expected when a single-bit quantizer is used. However (as pointed out in Chapter 1 (Section 1.4.4)), integrator nonlinearity results in significant harmonic distortion in single-bit CTDSMs, apart from a small increase in the in-band noise floor. This chapter involves the study of distortion observed in single-bit modulators. Breems *et al.* (1999) analyzed such effects when nonlinear single-stage active-RC integrators and an RZ feedback DAC are used in a single-bit modulator. We extend the analysis to other integrator topologies, considering both NRZ and RZ DACs. The nonlinear integrator models derived in the previous chapters will be used for this purpose.

4.1 Distortion with G_m-C and active-RC integrators with a single-stage transconductor

A. With NRZ DACs : A single-bit CTDSM having a nonlinear first integrator and an NRZ DAC can be represented as shown in Fig. 4.1(a). The effect of weak nonlinearity in the integrator is modeled as an additional input sequence, $z[n] = -\beta x^3[n]$ applied to

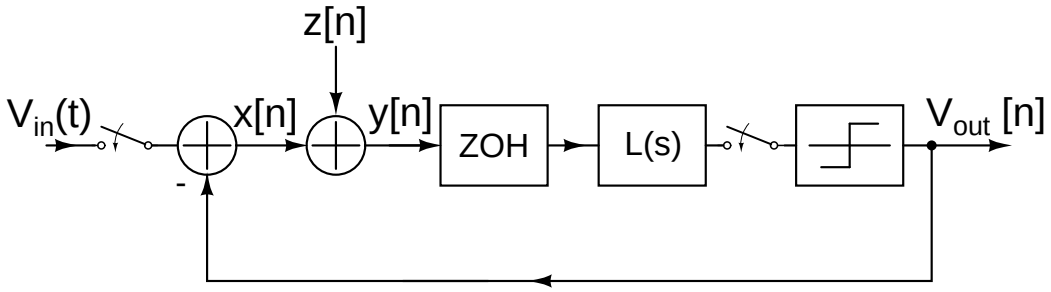


Figure 4.1: Model of a single-bit CTDSM with nonlinear G_m-C and active-RC integrators with a single-stage transconductor.

a linear modulator. Assuming the feedback DAC ($V_{out}[n]$) levels to be $\pm V_{ref}$, $y[n]$ is given as

$$\begin{aligned} y[n] &= x[n] - \beta x^3[n] && (4.1) \\ &= V_{in}[n] - V_{out}[n] - \beta V_{in}^3[n] + 3\beta V_{in}^2[n]V_{out}[n] \\ &\qquad -3\beta V_{in}[n]V_{out}^2[n] + \beta V_{out}^3[n] && (4.2) \\ &= (1 - 3\beta V_{ref}^2)V_{in}[n] - (1 - \beta V_{ref}^2)V_{out}[n] \\ &\qquad -\beta V_{in}^3[n] + 3\beta V_{in}^2[n]V_{out}[n] && (4.3) \end{aligned}$$

It can be seen that integrator nonlinearity in a single-bit CTDSM is not reflected in the DAC signal because of its binary nature ($\pm V_{ref}$). Assuming $V_{out} \approx V_{in}$ in the in-band region, we see that

$$y[n] \approx V_{in}[n] - V_{out}[n] + 2\beta V_{in}^3[n] \quad (4.4)$$

Thus, the nonlinear modulator can be considered to be a linear modulator with an additional input, $2\beta V_{in}^3$ as shown in Fig. 4.2. With A_{in} denoting the amplitude of the input signal, the third harmonic distortion observed at the modulator output is given by

$$HD_3 = \frac{\beta A_{in}^2}{2} \quad (4.5)$$

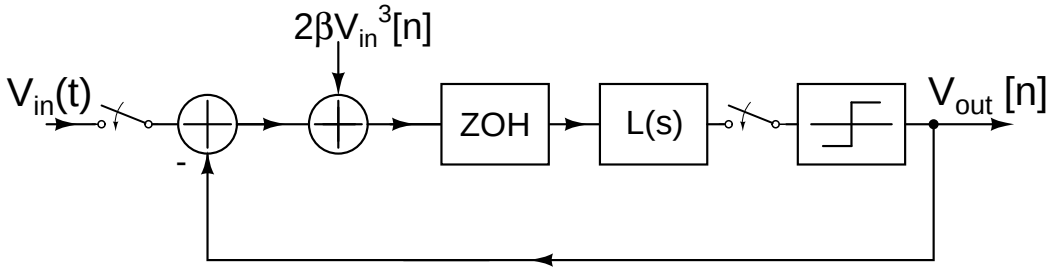


Figure 4.2: Equivalent model of the single-bit CTDSM with nonlinear single-stage integrators.

B. With RZ DACs : The model of a single-bit CTDSM with a nonlinear first integrator and an RZ DAC is shown in Fig. 4.3. The additional input sequences introduced due to integrator nonlinearity are defined as $z_1[n] = -4\beta x^3[n]$ and $z_2[n] = 6\beta x^2[n]V_{in}[n]$. Using the binary nature of the DAC signal and assuming $V_{out} \approx V_{in}$ in the signal band, the components (in $z_1[n]$ and $z_2[n]$) contributing to the third harmonic distortion are

$$z_1[n] \;=\; -4\beta x^3[n] \;\approx\; 8\beta V_{in}^3[n] \tag{4.6}$$

$$z_2[n] \;=\; 6\beta x^2[n]V_{in}[n] \;\approx\; -6\beta V_{in}^3[n] \tag{4.7}$$

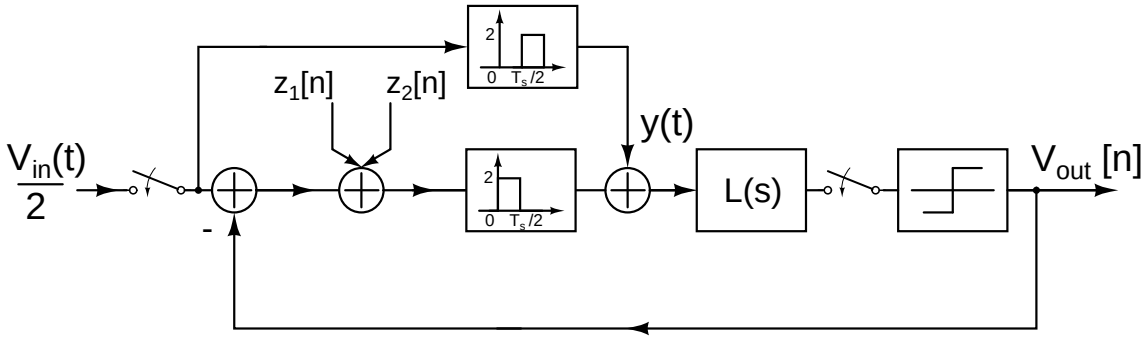


Figure 4.3: Model of a single-bit CTDSM with nonlinear single-stage integrator and an RZ DAC.

Thus, the effect of integrator nonlinearity can be modeled as an additional input sequence $2\beta V_{in}^3[n]$ to an otherwise linear modulator, as shown in Fig. 4.4.

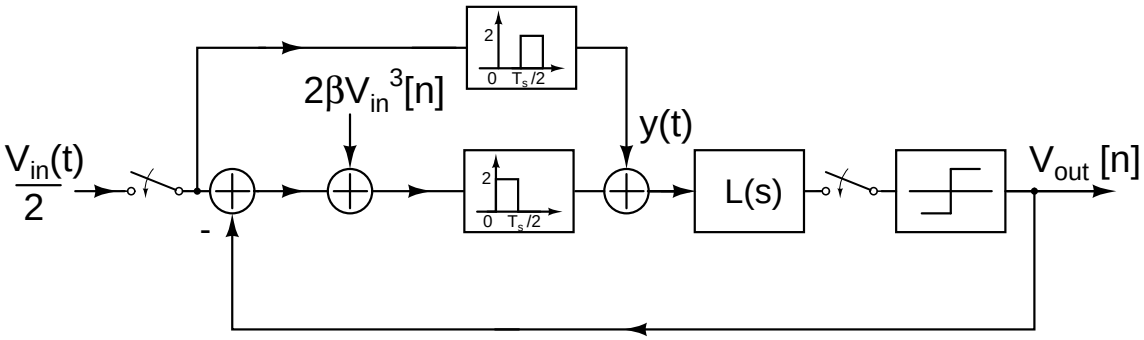


Figure 4.4: Equivalent model of the single-bit CTDSM with nonlinear single-stage integrators

The third harmonic distortion observed at the modulator output is therefore

$$HD_3 = \frac{\beta A_{in}^2}{2} \tag{4.8}$$

It can be recalled that the effect of integrator nonlinearity worsens in a multi-bit modulator when RZ DACs are employed. In contrast, it is seen that the performance of a single-bit CTDSM is the same with both NRZ and RZ DACs. Intuitively, this can be explained as follows. The feedback DAC signals do not get influenced by integrator nonlinearity because of their binary nature. Due to the mixing operation (or $(V_{in}[n] - V_{out}[n])^3$), the in-band signal component of $V_{out}[n]$ contributes to the third harmonic distortion. With the in-band signal component being independent of the DAC pulse shape, the performance can thus be expected to be identical.

While Breems *et al.* (1999) has quantified the distortion with a single-stage active-RC integrator (using a different approach), the models shown in Fig. 4.2 and Fig. 4.4 are applicable for both G_m-C and (single-stage) active-RC integrators, where

$$\beta = \frac{g_3}{4g_m}, \qquad \text{for a } \mathrm{G_m - C} \text{ integrator} \tag{4.9}$$

$$= \frac{2g_3}{g_m(2+g_mR)^3}, \qquad \text{for an active} - \mathrm{RC} \text{ integrator} \tag{4.10}$$

4.2 Distortion with active-RC integrators based on Miller and feedforward compensated opamps

A. With NRZ DACs : The discussion in Chapter 2 (Section 2.1.2) showed that active-RC integrators with two-stage opamps are better suited to meet the linearity requirements on the first integrator of CTDSMs. However, initial transients were found to degrade the performance of the modulator in these cases. The effect of nonlinearity in such integrators can be analyzed using the same model for the single-bit CTDSM, as with a single-stage active-RC integrator (shown in Fig. 4.5). However, the additional input sequence, $z[n]$ is now composed of four sequences, which results in

$$y[n] = (V_{in}[n] - V_{out}[n]) + \sum_{i=1}^{4} z_i[n] \tag{4.11}$$

where,

$$\begin{aligned} z_1[n] &= -\beta_1 x^3[n] \\ z_2[n] &= -\beta_2 x^2[n] x'[n] \\ z_3[n] &= -\beta_3 x[n] x'^2[n] \\ z_4[n] &= -\beta_4 x'^3[n] \end{aligned}$$

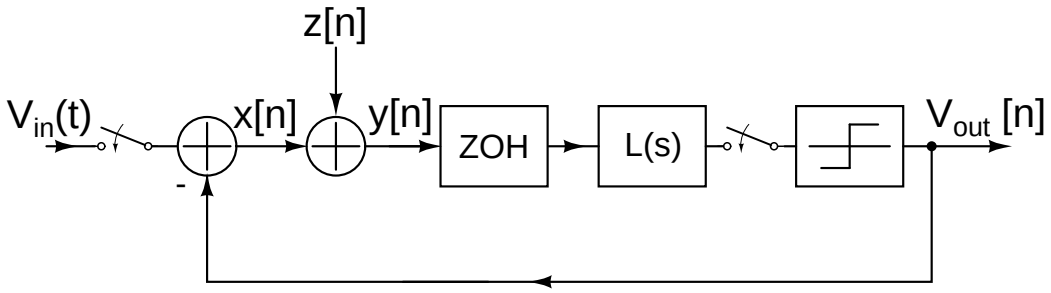


Figure 4.5: Model of a single-bit CTDSM with nonlinear two-stage opamp based active-RC integrators, where $z[n] = \sum z_i[n]$ (i=1,2,3,4).

where the nonlinear coefficients are defined (as in Chapter 3) as

$$\begin{aligned} \beta_1 &= \frac{2g_{31}}{g_{m1}(2+kg_{m1}R)^3} \\ \beta_i &= \frac{2g_{31}\alpha_i}{g_{m1}}, \qquad (\text{for i} = 2,3,4) \end{aligned}$$

and α_i denotes the time-average of the respective transient component, computed over one clock cycle. The components of $z[n]$ can be approximated in the in-band region as explained below.

The sequence, $x'[n] = x[n] - x[n-1]$ can be expressed as

$$x'[n] = \Big(V_{in}[n] - V_{in}[n-1]\Big) - \Big(V_{out}[n] - V_{out}[n-1]\Big) \tag{4.12}$$

With $\omega_{in} = 2\pi f_{in}/f_s$ denoting the input frequency (expressed in rad/sample), for slow varying input signals $V_{in}[n] - V_{in}[n-1] \approx \omega_{in}V_{in}[n]$. Hence,

$$x'[n] \approx \omega_{in}V_{in}[n] - \Big(V_{out}[n] - V_{out}[n-1]\Big) \tag{4.13}$$

Given the binary nature of the feedback DAC signal ($\pm V_{ref}$), higher powers of $x'[n]$ can be approximated as having the following components that can significantly contribute to harmonics of the input signal.

$$\begin{aligned} x'^{\,2}[n] &\approx -2\omega_{in}V_{in}[n]\Big(V_{out}[n] - V_{out}[n-1]\Big) \\ &\quad -2V_{out}[n]V_{out}[n-1] \end{aligned} \tag{4.14}$$

$$x'^{\,3}[n] \approx -6\omega_{in}V_{in}[n]V_{out}[n]V_{out}[n-1] \tag{4.15}$$

In a similar manner,

$$x^2[n] \approx V_{in}^2[n] - 2V_{in}[n]V_{out}[n] \tag{4.16}$$

$$x^3[n] \approx V_{in}^3[n] - 3V_{in}^2[n]V_{out}[n] \tag{4.17}$$

Using the above expressions, with $V_{out} \approx V_{in}$ in the signal band, the individual sequences ($z_i[n]$) can be simplified to

$$z_1[n] \approx 2\beta_1 V_{in}^3[n] \tag{4.18}$$

$$z_2[n] \approx 2\beta_2 V_{in}[n]V_{out}[n]V_{out}[n-1]\big) \tag{4.19}$$

$$z_3[n] \approx 2\beta_3 V_{in}[n]V_{out}[n]V_{out}[n-1] \tag{4.20}$$

$$z_4[n] \approx 6\beta_4\omega_{in} V_{in}[n]V_{out}[n]V_{out}[n-1] \tag{4.21}$$

The expression for $y[n]$ can now be given by

$$\begin{aligned} y[n] \approx\ & (V_{in}[n] - V_{out}[n]) + 2\beta_1 V_{in}^3[n] \\ & +2(\beta_2 + \beta_3 + 3\omega_{in}\beta_4)V_{in}[n]V_{out}[n]V_{out}[n-1] \end{aligned} \tag{4.22}$$

Comparing this expression with eq. 4.4, the presence of initial transients is seen to introduce a new term dependent on the product of the output signal ($V_{out}[n]$) with its 'delayed' version ($V_{out}[n-1]$). Simulations run on a third order (linear) audio modulator reveal two important observations concerning this component. Fig. 4.6 shows the in-band spectrum of $V_{in}^3[n]$ and $V_{in}[n]V_{out}[n]V_{out}[n-1]$, obtained through simulation. It can be seen that the third harmonic component present in $V_{in}[n]V_{out}[n]V_{out}[n-1]$ is higher than that found in $V_{in}^3[n]$. This can be attributed to the mixing of high-frequency (signal-

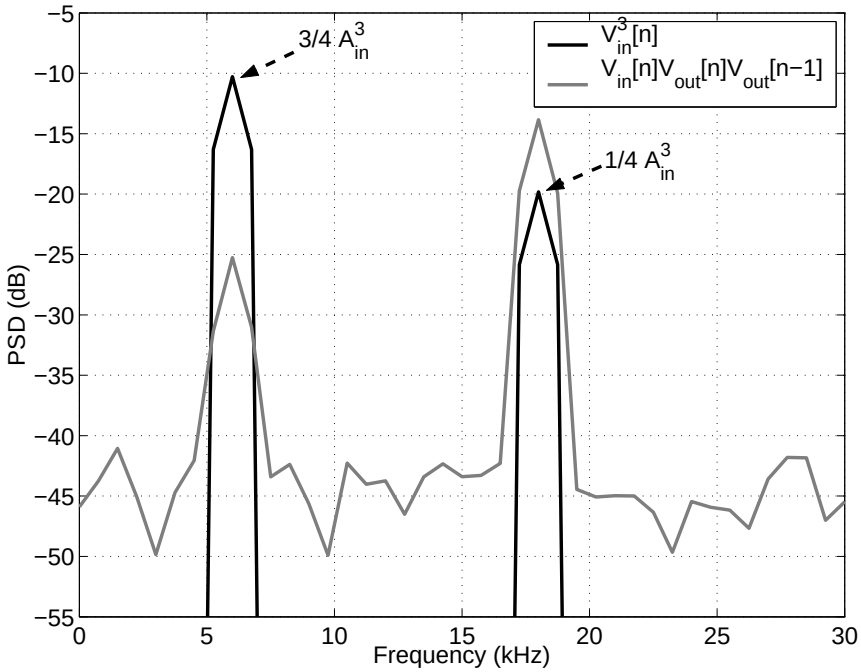


Figure 4.6: In-band PSD of $V_{in}^3[n]$ and $V_{in}[n]V_{out}[n]V_{out}[n-1]$, with 6 kHz being the input frequency.

dependent) tones that can be observed in the modulator output. This phenomenon can also be seen to result in demodulation of the out-of-band noise to low frequencies. This suggests that an increase in the in-band noise can be expected at high nonlinearities, apart from the increased third harmonic distortion with initial transients.

Let the factor γ_{in} denote the relative amplitude[1] of the third harmonic tone in the component $V_{in}[n]V_{out}[n]V_{out}[n-1]$, when compared to that present in $V_{in}^3[n]$. With A_{in} representing the input signal amplitude, the third harmonic distortion at the output of the modulator can be calculated from eq. 4.22 as

$$HD_3 \quad = \quad \Big(\beta_1 + \gamma_{in}(\beta_2 + \beta_3 + 3\omega_{in}\beta_4)\Big)\frac{A_{in}^2}{2} \tag{4.23}$$

[1]This can be determined through ideal modulator simulations, for the given input amplitude (normalized to the DAC reference V_{ref}).

In summary, similar to the multi-bit modulators, the effect of initial transients is seen to degrade the performance of single-bit CTDSMs. The mixing between $V_{out}[n]$ and $V_{out}[n-1]$ is observed to lead to an increased third harmonic distortion in these cases. Such mixing operations occur because of the interference between two consecutive output samples of the modulator, when an NRZ DAC is used. Since modulators with RZ DAC are free from any such interferences, it would be interesting to investigate if the effects of the initial transients are negligible.

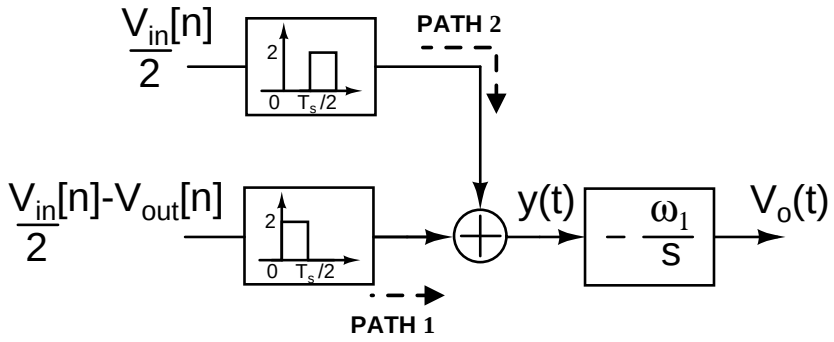


Figure 4.7: Model for the input stage of a CTDSM with an RZ feedback DAC and a linear first integrator.

B. With RZ DACs : The input stage of a CTDSM having an RZ feedback DAC and a linear first integrator is shown in Fig. 4.7. As in the previous chapter (Section 2.2.1), two paths (1 and 2) with complementary RZ DACs have been used to model the input signals of the integrator. The input to the integrator ($y(t)$) is given as

$$y(t) = V_{in}[n] - 2V_{out}[n]\,, \quad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+0.5)\mathrm{T_s} \tag{4.24}$$

$$= V_{in}[n]\,, \quad (\mathrm{n}+0.5)\mathrm{T_s} \leq \mathrm{t} < \mathrm{nT_s} \tag{4.25}$$

Similar to the case with an NRZ feedback DAC, the input to the integrator can be considered as a series of step-like signals. However, with the RZ DAC, the steps change for every $T_s/2$ interval. During the first half (or phase) of the clock cycle, the input to

the integrator comprises

- a pulse of amplitude $V_{in}[n] - 2V_{out}[n]$ $(= 2x_{rz}[n])$
- along with a transition of height $V_{in}[n] - V_{in}[n-1] - 2V_{out}[n]$

For simplicity, assuming a high OSR, $V_{in}[n] \approx V_{in}[n-1]$. Assuming $w_n(t)$ to denote $w(t - nT_s)$, the linear response at v_d can be expressed as

$$v_{d,lin}(t) \approx 2G_0 x_{rz}[n] - 2V_{out}[n]w_n(t)\,, \quad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+0.5)\mathrm{T_s} \tag{4.26}$$

From the knowledge of the linear response, the injected current ($i_{inj} = g_{31}v_{d,lin}^3$) for $\mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+0.5)\mathrm{T_s}$ can be determined as

$$\begin{aligned} i_{inj}(t) &= 8g_{31}G_0^3x_{rz}^3[n] - 24g_{31}G_0^2x_{rz}^2[n]V_{out}[n]w_n(t) \\ &\quad + 24g_{31}G_0x_{rz}[n]V_{out}^2[n]w_n^2(t) - 8g_{31}V_{out}^3[n]w_n^3(t) \end{aligned} \tag{4.27}$$

With $V_{out}^2[n] = V_{ref}^2$, only the first two terms in the above expression can contribute to third harmonic distortion. Thus, the injected current can be simplified to

$$i_{inj}(t) \approx 8g_{31}G_0^3x_{rz}^3[n] - 24g_{31}G_0^2x_{rz}^2[n]V_{out}[n]w_n(t) \tag{4.28}$$

$$\approx 8g_{31}G_0^3x_{rz}^3[n] - 6g_{31}G_0^2V_{in}^2[n]V_{out}[n]w_n(t)\,,\ \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+0.5)\mathrm{T_s} \tag{4.29}$$

During the second half of the clock cycle ($(\mathrm{n}+0.5)\mathrm{T_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s}$), the input to the integrator consists of a pulse of amplitude $V_{in}[n]$ with a transition of height $2V_{out}[n]$ at every $t = (n+0.5)T_s$. The injected current during this phase is

$$i_{inj}(t) = g_{31}\Big(G_0V_{in}[n] + 2V_{out}[n]w_{n+0.5}(t)\Big)^3 \tag{4.30}$$

$$\begin{aligned} &= g_{31}G_0^3V_{in}^3[n] + 6g_{31}G_0^2V_{in}^2[n]V_{out}[n]w_{n+0.5}(t) \\ &\quad + 12g_{31}G_0V_{in}[n]V_{out}^2[n]w_{n+0.5}^2(t) + 8g_{31}V_{out}^3[n]w_{n+0.5}^3(t) \end{aligned} \tag{4.31}$$

It is apparent that only the first two terms are significant, thus resulting in

$$i_{inj}(t) \approx g_{31}G_0^3V_{in}^3[n] + 6g_{31}G_0^2V_{in}^2[n]V_{out}[n]w_{n+0.5}(t) \tag{4.32}$$

for $(\mathrm{n}+0.5)\mathrm{T_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s}$. From the expression for the injected current during the two phases of the clock cycle (eq. 4.29 and 4.32), it is clear that the currents that are dependent on the initial transient ($w(t)$) will get averaged out in a given clock cycle. In essence, the injected current is seen to be purely determined by G_0, similar to the case of single-stage active-RC integrators. The third harmonic distortion observed at the modulator output can easily be shown to be

$$HD_3 = \frac{\beta_1 A_{in}^2}{2} \tag{4.33}$$

where $\beta_1 = 2g_{31}/(g_{m1}(2+kg_{m1}R)^3)$. Such a performance advantage observed with RZ DACs can be attributed to two reasons. Firstly, equal and opposite transitions in the DAC signal can be thought as nullifying the error charges injected in each clock cycle. In essence, transients in response to the odd powers of the DAC signal ($V_{out}[n], V_{out}^3[n]$) effectively inject zero error charge in each cycle. Secondly, thanks to the binary nature of the DAC signal, the charges injected due to the components containing even powers of V_{out} (say $V_{in}[n]V_{out}^2[n]$), also do not lead to harmonic distortions.

However, such an advantage with RZ DACs should be appreciated with caution owing to the increased jitter and slew-rate requirements[2]. In addition, due to the presence of initial transients, the transconductors can be pushed into stronger regions of nonlinearity. In such cases, the third harmonic distortion will worsen than predicted in the above analysis.

[2]The peak input to the loop filter is $V_{in,max}+2V_{ref}$ as opposed to $V_{in,max}+V_{ref}$ for the NRZ case.

4.3 Simulation results

A third order single-bit modulator was chosen as a test vehicle to verify the results derived in the previous sections. The NTF has an OBG of 1.5 with complex in-band zeros. With an OSR of 128, a peak SNR of 110 dB can be achieved with this NTF in a 24 kHz bandwidth. The feedback DAC reference voltage (V_{ref}) was assumed to be 1.8 V and the first integrating resistor was chosen to be R = 150 kΩ. The effect of nonlinearity was observed for different topologies in the first integrator, using the same opamps as in our simulations with multi-bit modulator. In all the simulations, distortion was measured by applying an input tone of 1 V amplitude at a frequency of 6 kHz.

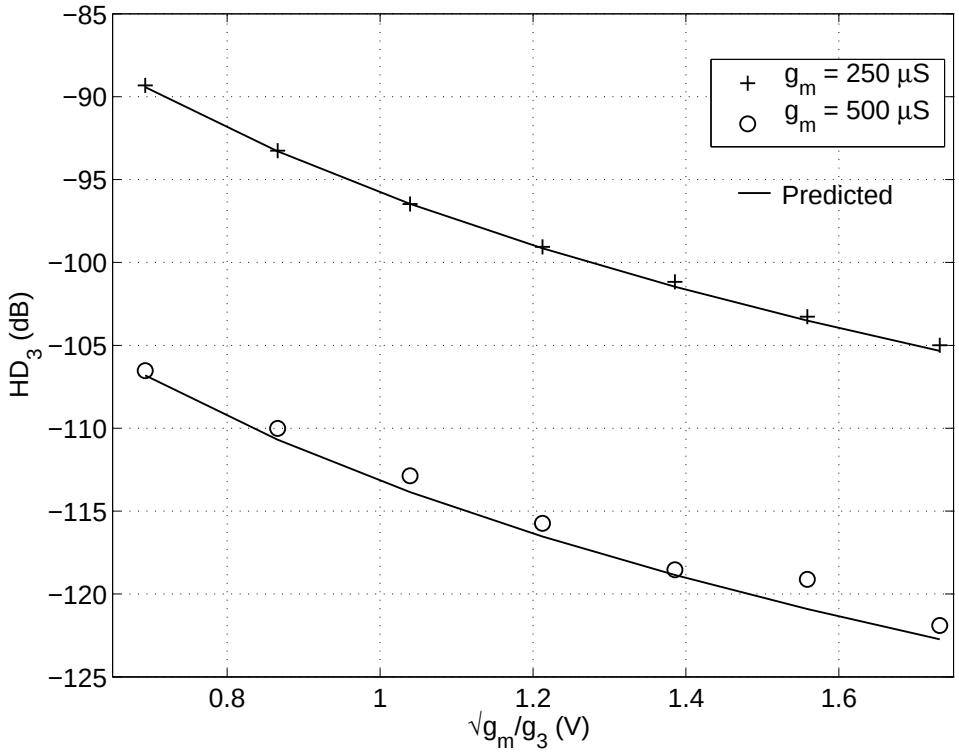


Figure 4.8: Variation of HD_3 with a single-stage transconductor based active-RC integrator for two values of g_m.

A. Performance with an NRZ DAC : Fig. 4.8 shows the variation[3] in HD_3 observed with a single-stage active-RC integrator for two values of the transconductance, g_m. The first integrating capacitor was fixed at 5.77 pF in these simulations. The predicted distortion is seen to closely match the simulated results. Fig. 4.9 shows the observed and predicted variation in HD_3, obtained with the Miller opamp based active-RC integrator. The variation is plotted for different values of the integrating capacitor, with Table 4.1 showing the respective β_i/β_1 used in calculations. The factor γ_{in} was found to be 2.24 through ideal modulator simulations for the given input amplitude of 1 V.

Opamp topology/C	$\boldsymbol{\beta_2/\beta_1}$	$\boldsymbol{\beta_3/\beta_1}$	$\boldsymbol{\beta_4/\beta_1}$
Miller/5.77 pF	2.15	17.1	41.5
Miller/2.88 pF	1.07	4.3	5.2
Miller/8.65 pF	3.2	37.75	135.8
Feedforward/5.77 pF	4.35	73.25	507.9
Feedforward/2.88 pF	2.48	23.9	94.45
Feedforward/8.65 pF	5.86	132.9	1240

Table 4.1: Relative nonlinear coefficients of various integrator configurations used in the single-bit modulator with NRZ DAC. γ_{in} = 2.24 for 1 V input amplitude.

Similar tests were run with the feedforward based integrator and Fig. 4.10 shows the observed and predicted variation in HD_3 for different nonlinearities in the input transconductor. As expected, the performance can be seen to improve for increasing values of the integrating capacitor, though the improvement is only directly proportional to C. As can be inferred from Table 4.1, this is due to the significance of the initial transients and in particular, of the factor β_3 (which varies approximately as $1/C$).

Another effect of the initial transients that can be observed to affect the performance of single-bit CTDSMs is shown in Fig. 4.11. In addition to the harmonic distortion, an

[3] Since the harmonic distortion is purely determined by the input signal amplitude and the nonlinearity of the integrator, the x-axis is provided in absolute terms of $\sqrt{g_m/g_3}$.

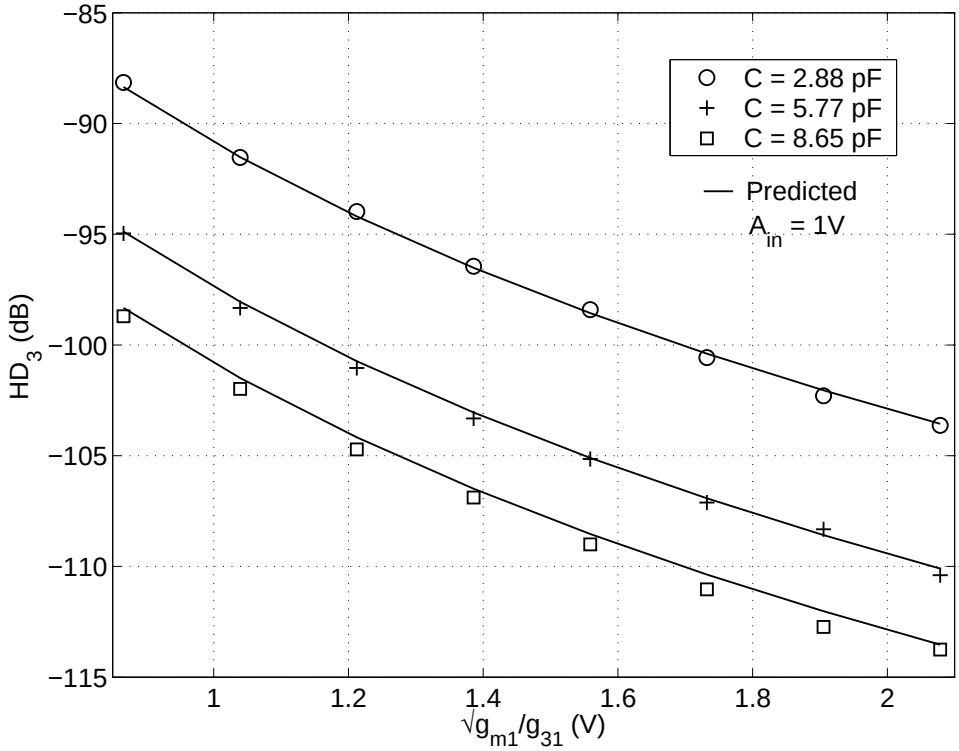


Figure 4.9: Variation of HD_3 with Miller compensated opamp due to nonlinearity in the input transconductor for different integrating capacitors.

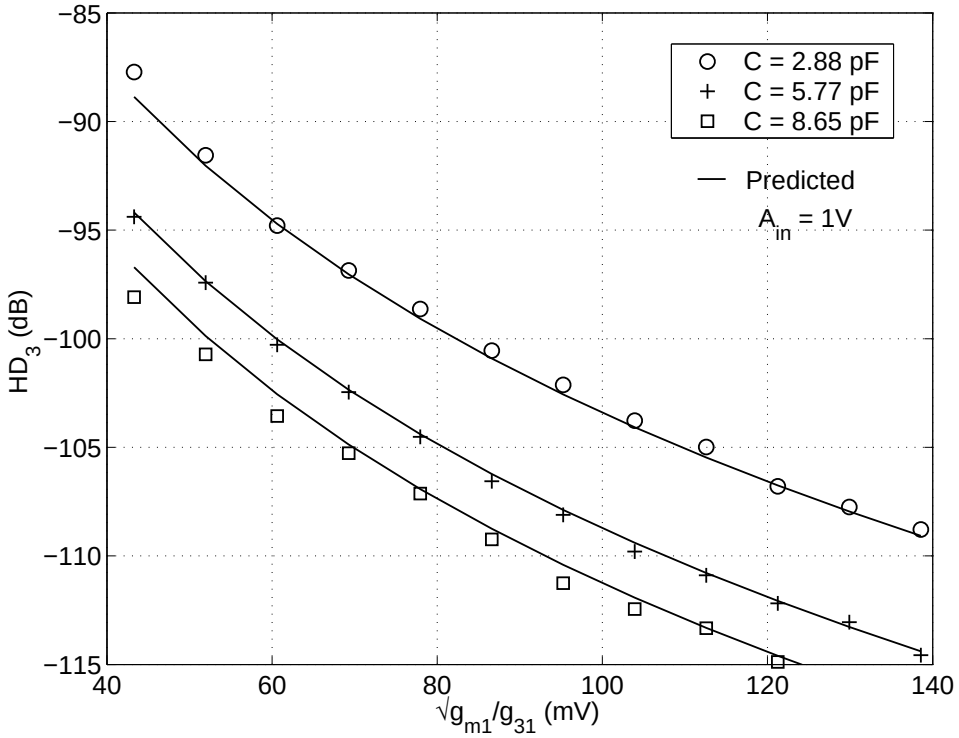


Figure 4.10: Variation of HD_3 with feedforward compensated opamp due to nonlinearity in the input transconductor for different integrating capacitors.

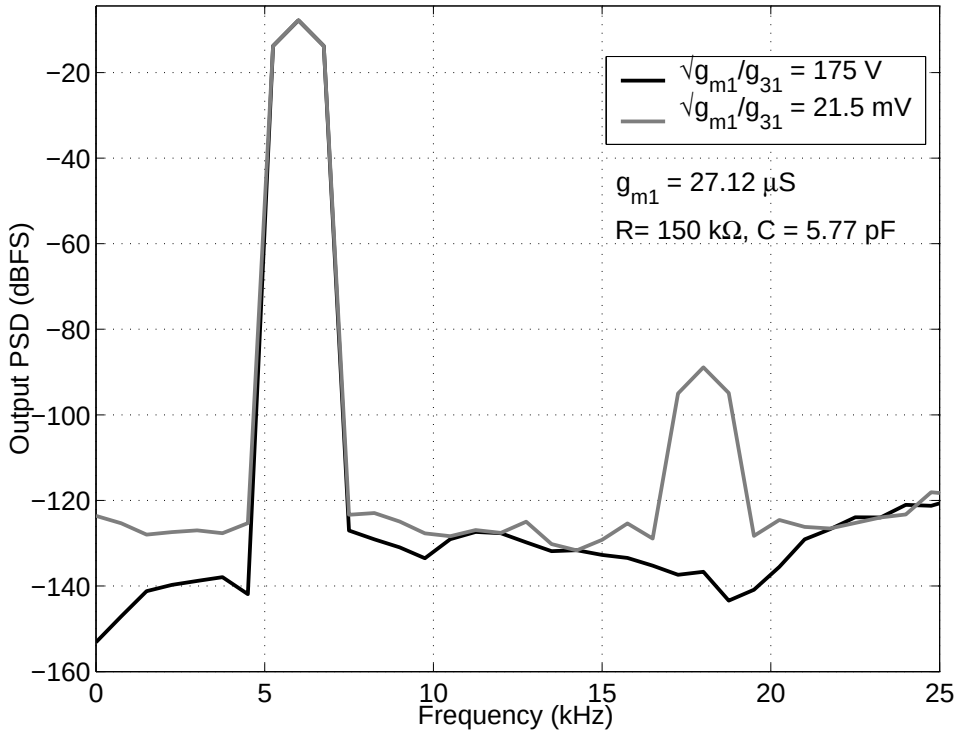


Figure 4.11: In-band PSD of the modulator output showing the increase in noise floor at high nonlinearities

increase in the in-band noise can be observed at very high nonlinearities ($\sqrt{g_{m1}/g_{31}}$ = 21.5 mV). This can be attributed to the demodulation of the out-of-band noise to the in-band region, arising out of the mixing operation between $V_{out}[n]$ and $V_{out}[n-1]$. Further, the dependency of the harmonic distortion with input frequency (as given in eq. 4.23) was confirmed for two different integrating capacitors as shown in Fig. 4.12. It can be seen that the distortion increases with input frequency, with the variation being higher with C = 8.65 pF than with C = 5.77 pF. Such a trend is due to the higher value of β_4/β_1 (see Table 4.1) for the case of C = 8.65 pF.

B. Performance with an RZ DAC : To demonstrate the accuracy of the models with an RZ DAC, the modulator was simulated with the same nonlinear integrators as in the NRZ case. Fig. 4.13 compares the observed and predicted variation in HD_3 obtained with the Miller opamp based active-RC integrator, with an integrating capacitor

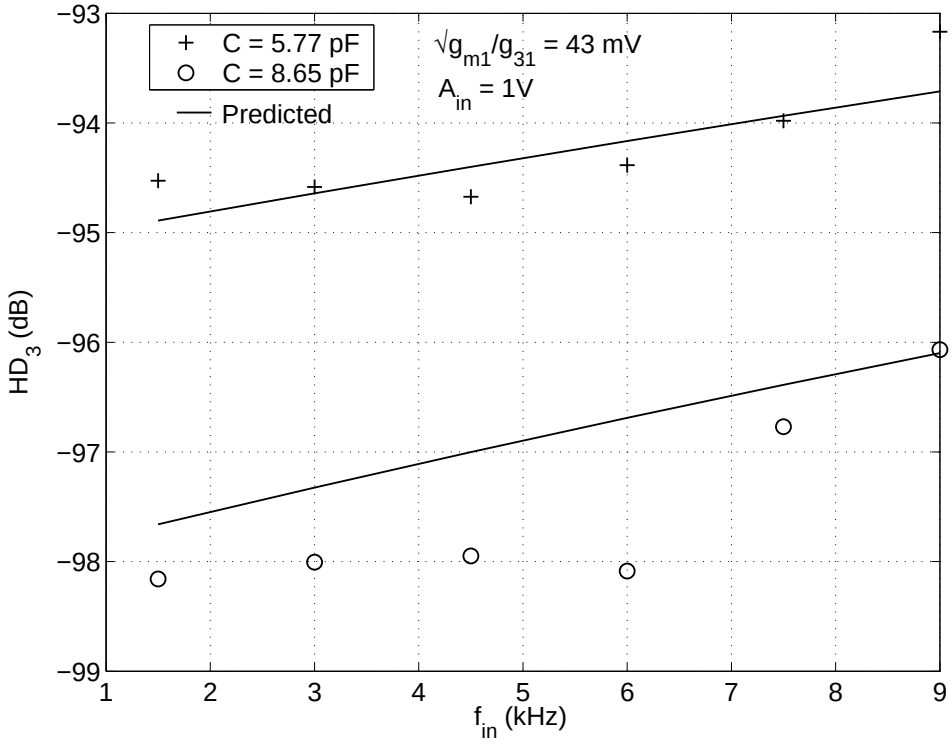


Figure 4.12: Variation of HD_3 (with feedforward compensated opamp) as a function of the input frequency for different values of integrating capacitors.

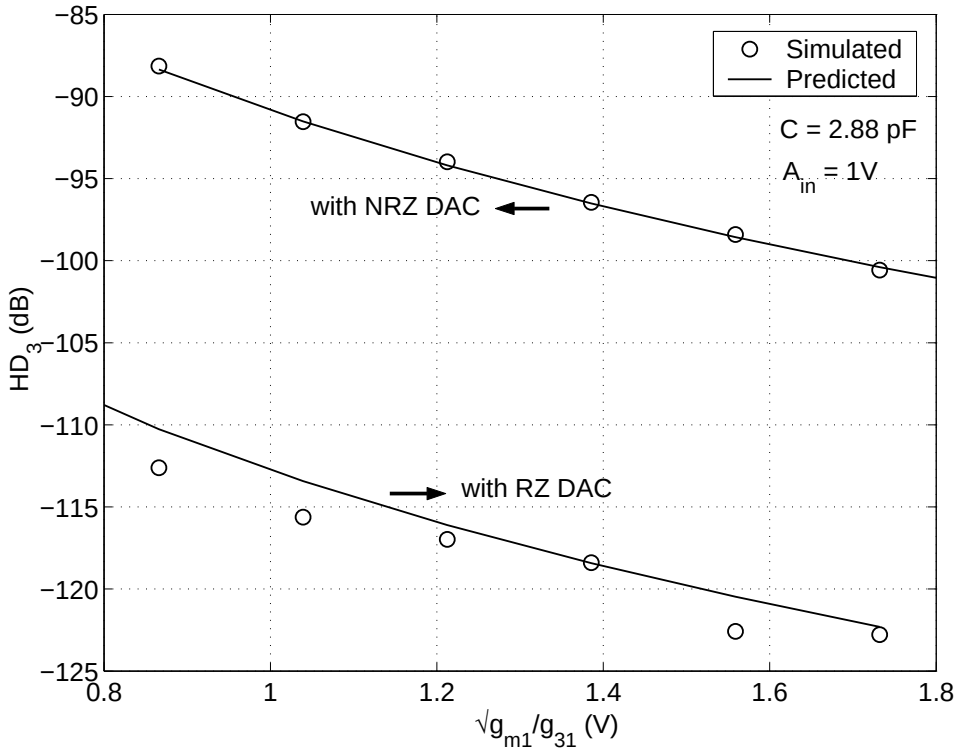


Figure 4.13: Variation of HD_3 with a Miller compensated opamp as function of nonlinearity in the input transconductor, in modulators with NRZ and RZ DAC.

of 2.88 pF. There is good matching between the results from theory and simulation. Also, observe that the performance with an RZ DAC is better than with an NRZ DAC, even for an integrating capacitor as low as 2.88 pF[4]. A similar performance improvement can be seen in Fig. 4.14, where the in-band PSD of both the NRZ and RZ modulator outputs has been plotted for the case of the feedforward opamp based active-RC integrator. With an integrating capacitor of 2.88 pF and the nonlinearity of the input transconductor fixed at $\sqrt{g_{m1}/g_{31}}$ = 87 mV, the harmonic distortion observed with the NRZ modulator is -100 dB. In comparison, thanks to the large $k \approx C/c_p$ achieved with a feedforward architecture for the opamp, the RZ modulator is free of third harmonic distortion.

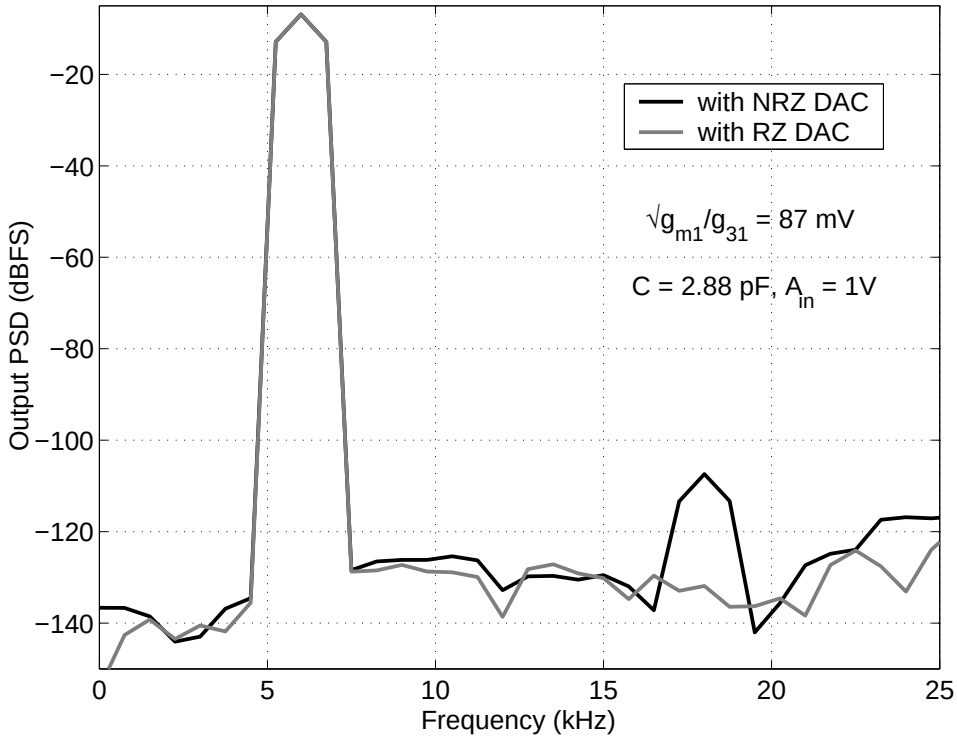


Figure 4.14: In-band PSD of the single-bit CTDSM employing a nonlinear feedforward opamp based active-RC integrator, comparing the performance with NRZ and RZ DAC

[4]While the integrating capacitor can in principle be reduced further, integrator swings exceeded the supply limits for lower values of C. For the given single-bit CTDSM with RZ DAC, an integrating capacitor of 2.88 pF limits the swing to 2.4 V (peak-to-peak differential), which is two-third of the supply or the full-scale range.

4.4 Severity of integrator nonlinearity in single-bit and multi-bit modulators and discussion

The knowledge of the effect of integrator nonlinearity obtained for both single-bit and multi-bit modulators can now be used to compare the severity of integrator nonlinearity in each of these modulators. The comparison can be made based on the performance with feedforward opamp based active-RC integrator, since feedforward opamps prove to be the best choice for designing low power, high resolution CTDSMs[5].

Referring to Fig. 3.11 in the previous chapter, a 4-bit modulator employing a feedforward opamp based active-RC integrator can be observed to provide a performance equivalent[6] of 110 dB, at nonlinearity levels as low as $\sqrt{g_{m1}/g_{31}}$ = 0.1 LSB. This is achieved with an integrating capacitor of about 1 pF. In contrast, (as seen from Fig. 4.10) such a performance is possible with the single-bit modulator only at relatively weaker levels of nonlinearity (or larger $\sqrt{g_{m1}/g_{31}}$) and with smaller integrator bandwidths (or larger RC).

This confirms the general intuition that opamp nonlinearity becomes severe as the quantizer bits are reduced. However, a key point to note in case of a single-bit CTDSM is that the performance is degraded through harmonic distortions in the modulator output. Since the distortions worsen with increasing signal power, this reduces the maximum Signal-to-Noise-plus-Distortion Ratio (SNDR) that can be achieved. This will significantly affect the dynamic range of the modulator unless corrective measures are adopted.

[5] either single-bit modulator (see Fig. 4.9 and 4.10) or multi-bit modulator.

[6] The performance of the multi-bit modulator can be compared with the ideal SNR obtained with the given single-bit CTDSM.

To briefly summarize the possible measures, an apparent way of improving the performance is to increase the resistance, R. Since the thermal noise of the modulator increases in this case, there is a limit to the maximum value of the resistor tolerable. As an alternate option, the integrating capacitor (C) can also be increased. However, unlike with the resistor, an increase in its value by a factor of 2 does not improve the performance by 18 dB. This is due to the fact that the initial transients ($w(t)$) are not reduced by the increase in the integrating capacitor. As discussed in the previous section, since HD_3 is primarily determined by β_3 ($\propto 1/C$), the performance is found to improve only in steps of 6 dB in such cases.

To elaborate on the significance of the initial transients, consider the case where C = 5.77 pF. With the feedforward opamp topology, a very high value for k (= 429) can be obtained with the given value of the integrating capacitor. For $\sqrt{g_{m1}/g_{31}}$ = 87 mV, β_1 ($= 2g_{31}/g_{m1}(2 + kg_{m1}R)^3$) can be calculated to be about 50 nV^{-2}. If the effect of initial transients are neglected, the HD_3 can thus be expected to get lowered to -152 dB. However, as seen from Table 4.1 and Fig. 4.10, the presence of initial transients (or β_3) has led to a harmonic distortion as high as -106 dB. In such situations, the performance of the modulator can only be enhanced by

(a) improving the opamp dynamics (g_{m3}, g_{m2}), to reduce the initial transients or
(b) by increasing the linearity of the transconductors.

Unfortunately, either of these solutions will lead to an increased power consumption of the first integrator and thereby the power dissipated in the loop filter. For this reason, designing a single-bit CTDSM with low power can be challenging, thus making multi-bit modulators more preferable for their reduced linearity requirements.

CHAPTER 5

MITIGATION OF NONLINEARITIES USING THE ASSISTED OPAMP TECHNIQUE

The discussion in the previous chapter indicated that the design of the first integrator with adequate linearity and low power can pose significant challenges while designing high performance single-bit CTDSMs. The effect of initial transients in the two-stage active-RC integrators[1] was shown to limit the performance of modulators with an NRZ feedback DAC. In this chapter, a power efficient solution that addresses this problem will be introduced.

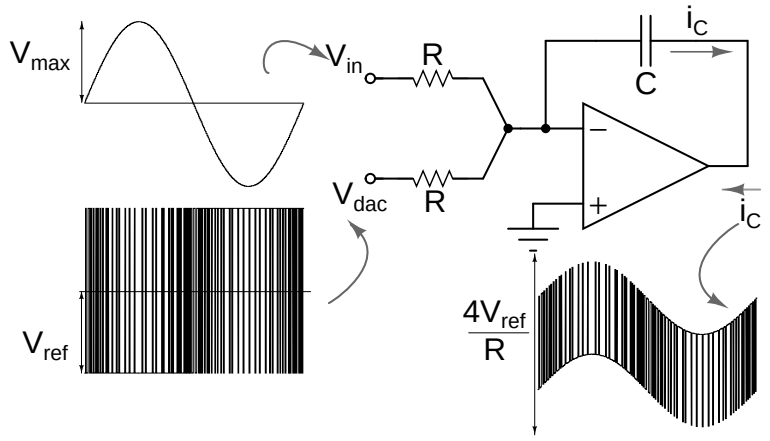


Figure 5.1: Schematic and waveforms in a conventional integrator

To motivate the proposed solution, consider the first integrator in a conventional single-bit modulator, as shown in Fig. 5.1. For simplicity, a single-ended circuit is assumed. The input is a sinusoid, with an amplitude V_{max}. The feedback DAC waveform takes on one of two values $\pm V_{ref}$. The current i_C flowing through the integrating capacitor, can be seen to have a peak-to-peak value of almost $4V_{ref}/R$. This current can be expected

[1] which are preferred to other integrator topologies for better linearity

to mostly consist of high frequency contents, due to the voltage steps introduced by the feedback DAC. If the opamp is ideal, the virtual ground voltage remains at zero and there is no degradation in the modulator performance.

A real opamp, however, cannot respond fast enough to the sharp steps of the DAC feedback voltage. This effect manifests as jumps (or initial transients) in the virtual ground potential whenever the feedback DAC switches. If the opamp is nonlinear, these transients significantly degrade the performance of the modulator. The conventional way of addressing this issue is to bias the opamps with large currents, so that the bandwidth and/or the linearity of the opamp are enhanced. Observe that the need for the opamp to source or sink large currents "instantaneously" decides the bandwidth required in the opamp. Instead of improving the bandwidth with increased power, the following solution can thus be suggested.

The current demanded of the opamp is known exactly, since one has access to V_{in} and V_{dac}. This information can be used to "assist" the opamp in discharging its function, as shown in Fig. 5.2. A transconductor with $g_m = 1/R$ and a replica current steering DAC with outputs $\pm V_{ref}/R$ (referred to as the assistant) pull current out of the opamp - this way, the opamp does not have to supply any current.

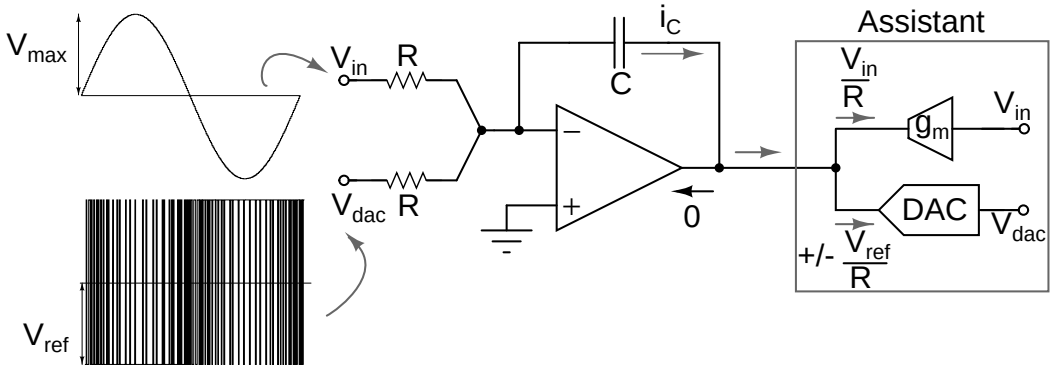


Figure 5.2: Assisted opamp integrator

Since the steps of DAC current are now supplied by the assistant, the excursions at the virtual ground node are greatly reduced. This results in low distortion operation of the modulator. In effect, the only requirement on the opamp is to supply the load currents and any incidental differences between the integrator input current and the assistant current. Therefore, the power requirements of the opamp can be significantly relaxed. We call this technique for designing linear, low power integrators as the 'assisted opamp' technique. The rest of the chapter provides a detailed description of this technique.

5.1 Basics of opamp assistance

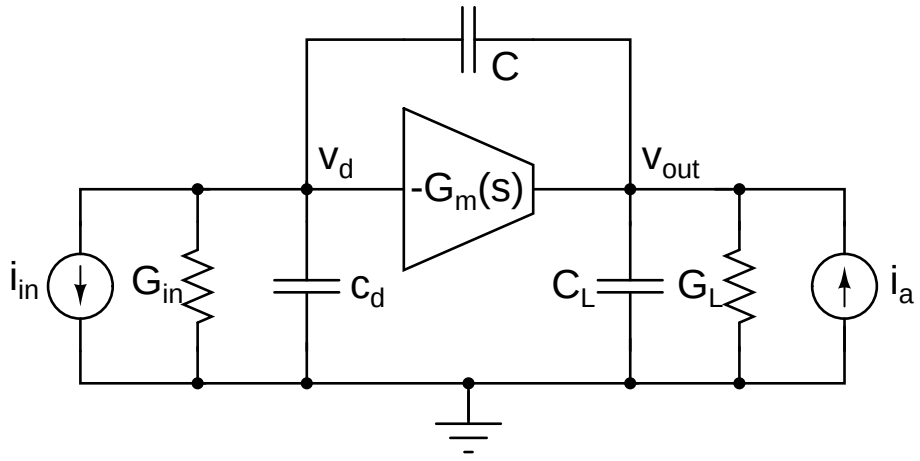


Figure 5.3: Equivalent circuit of an assisted opamp integrator

The fundamentals of opamp assistance can be understood using the equivalent circuit shown in Fig. 5.3, where some of the practical nonidealities are shown. i_{in}, G_{in} denote the Norton equivalent of the input and DAC circuits, with i_a denoting the assistant current. c_d is the capacitance at the virtual ground node. C is the integrating capacitor, while C_L and G_L are the total load capacitance and conductance that the opamp needs to drive. The opamp itself is modeled as a frequency dependent transconductor[2], $G_m(s)$. The Laplace transform of the response at the virtual ground node, v_d is

[2]This is an accurate representation of single-stage and two-stage feedforward compensated opamps.

given by

$$V_d(s) = \frac{-I_{in}(s)\big(s(C+C_L)+G_L\big) + I_a(s)sC}{G_{in}G_L + sC\big(G_m(s) + G_L(1+\frac{c_d}{C}) + G_{in}(1+\frac{C_L}{C})\big) + s^2(c_dC + c_dC_L + CC_L)} \tag{5.1}$$

From the above equation, we can see that v_d will not swing if

$$I_a(s) = \left(1+\frac{C_L}{C}\right) I_{in}(s) + \frac{G_L}{sC} I_{in}(s) \tag{5.2}$$

This indicates that the assistant current must contain proportional and integral versions of the input current to ensure zero swing at v_d. Note that in a CTDSM, the current through G_L is usually small, in comparison with the current flowing through the integrating capacitor[3]. Alternately, if subsequent integrators and the summation of their outputs are implemented using G_m-C techniques, the first integrator's resistive load (G_L) becomes negligible. Assuming $G_L = 0$, the swing at the virtual ground node can thus be canceled with an assistant current given by

$$I_a(s) \approx \left(1+\frac{C_L}{C}\right) I_{in}(s) \tag{5.3}$$

The above equation indicates that the cancellation at v_d *does not* depend on c_d, G_{in} or more importantly, on the shape of i_{in}. This implies that the idea of opamp assistance is not restricted to be used with NRZ DACs.

Fig. 5.4(a) and (b) illustrate how opamp assistance can be used for integrators based on single-stage and two-stage feedforward compensated opamps. As discussed before, the assistant current is obtained using a transconductor (that provides the input component of the integrator current) and a current steering DAC (which provides the DAC

[3]Since subsequent integrators are impedance and node scaled to reduce power dissipation.

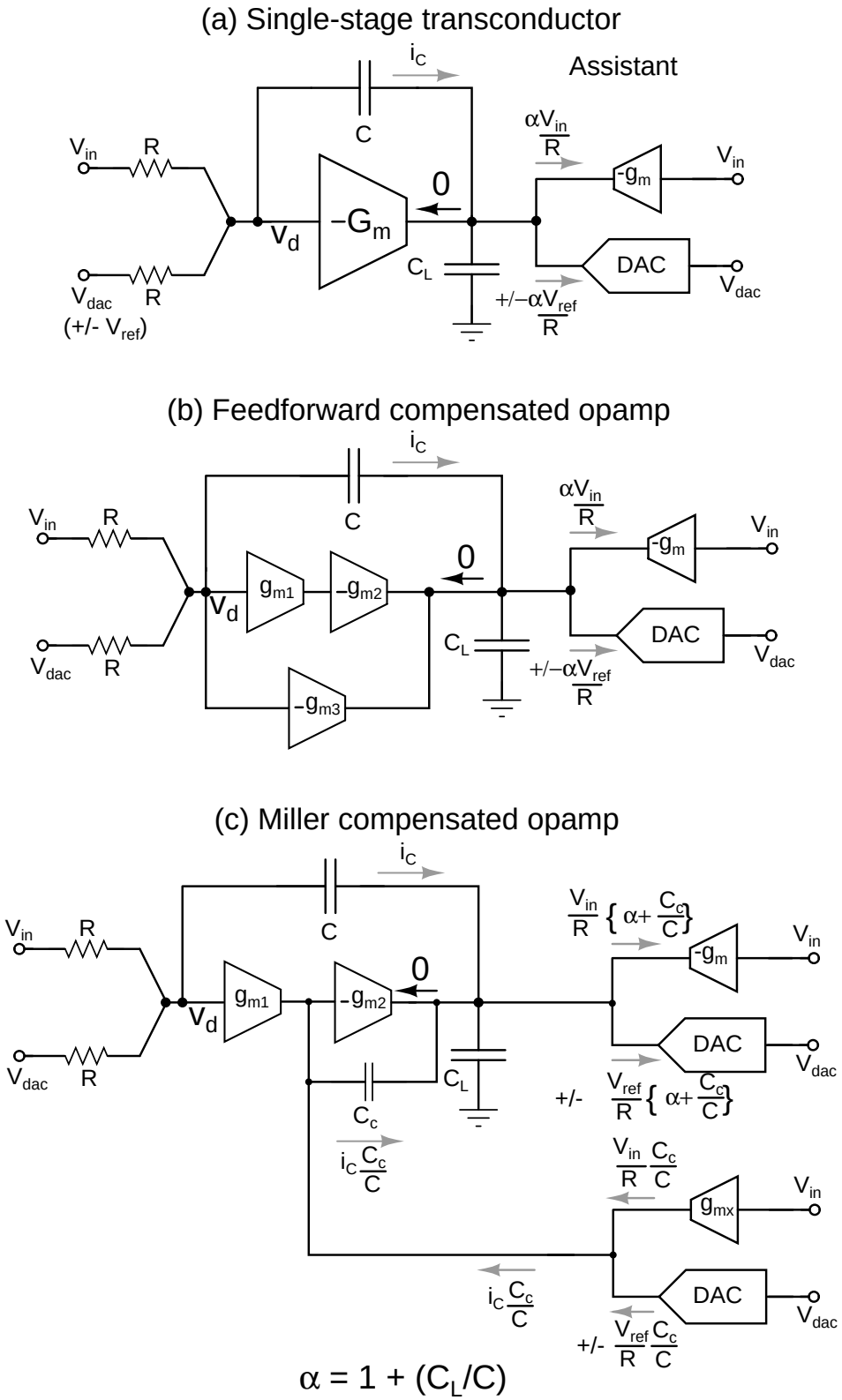


Figure 5.4: Implementation of the assisted opamp integrator for several commonly used opamp topologies - (a) A single stage transconductor (b) A two-stage feedforward compensated opamp and (c) A Miller compensated opamp ($\alpha = 1 + C_L/C$)

component of the integrator current). A Miller compensated two-stage opamp can also be "assisted", as shown in Fig. 5.4(c). Two assistants are needed in this case - apart from the one that injects current into the output, another assistant is needed to remove the current $i_C C_c/C$ that flows through the compensating capacitor. As a result, the input transconductor (g_{m1}) does not need to sink any current, resulting in $v_d = 0$.

Having understood the usefulness of the assisted opamp technique, a more quantitative analysis of integrator nonlinearity can now be obtained using the methodology discussed in Chapter 2 (Section 2.1.2). Since the above discussion can be directly extended to analyze single-stage integrators, attention will be given to active-RC integrators with two-stage opamps.

5.2 Analysis of nonlinearity in two-stage opamp assisted active-RC integrators

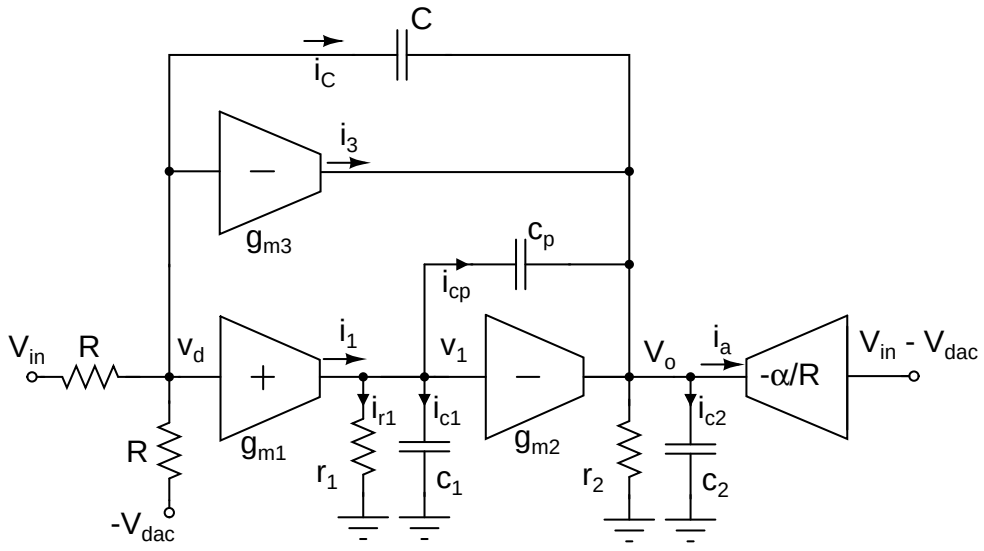


Figure 5.5: Assisted active-RC integrator with feedforward compensated opamp

Consider an active-RC integrator having a feedforward compensated opamp, that is provided with opamp assistance as shown in Fig. 5.5. The assistant transconductor and the DAC have been modeled together as a single transconductor pulling a current

$\alpha(V_{in} - V_{dac})/R$. Let α be defined as

$$\alpha = m_1 \left(1 + \frac{c_2}{C}\right) \tag{5.4}$$

where the factor m_1 can model any variation in the assistant current from the expected value (ideally $m_1 = 1$). Denoting the assistant current as i_a, the step-like voltage component at v_1 is

$$v_{1,cap} = \frac{(i_C + i_{cp} - i_{c2} - i_3)}{g_{m2}} - \frac{i_a}{g_{m2}} \tag{5.5}$$

$$\approx \frac{i_C(1 + c_p/C + c_2/C) - i_3}{g_{m2}} - \frac{\alpha(V_{in} - V_{dac})}{g_{m2}R} \tag{5.6}$$

Assuming $i_C \approx (V_{in} - V_{dac})/R$, we have

$$v_{1,cap} \approx \frac{i_C(1 - \alpha + c_p/C + c_2/C) - i_3}{g_{m2}} \tag{5.7}$$

The above modified expression for $v_{1,cap}$ thus leads to a modified i_{r1} given by

$$i_{r1} \approx \frac{v_{1,cap}}{r_1} \tag{5.8}$$

$$= \frac{i_C(1 - \alpha + c_p/C + c_2/C) - i_3}{g_{m2}r_1} \tag{5.9}$$

The input transconductor's current $i_1 = i_{cp} + i_{c1} + i_{r1}$ is given by

$$i_1 \approx i_C \left(\frac{c_p}{C}(1 + \frac{1}{g_{m2}r_2}) + \frac{c_1}{C}\frac{1}{g_{m2}r_2} + \frac{1 - \alpha + c_p/C + c_2/C}{g_{m2}r_1}\right) - \frac{i_3}{g_{m2}r_1} \tag{5.10}$$

Comparing the above expression with that in eq. 2.48, the linear response at the virtual

ground node can be shown to be

$$v_{d,lin} \approx \frac{V_{in}-V_{dac}}{2+k\Big(g_{m1}+\dfrac{g_{m3}}{g_{m2}r_1}\Big)R} \approx \frac{V_{in}-V_{dac}}{2+kg_{m1}R} \tag{5.11}$$

$$\text{where,} \quad k = \frac{C}{c_p+\dfrac{c_1+c_p}{g_{m2}r_2}+\dfrac{(C+c_2)(1-m_1)+c_p}{g_{m2}r_1}} \tag{5.12}$$

The fact that the assistant now provides the current i_C and i_{c2} can clearly be seen to modify the expression for k and thereby the swing at the virtual ground node. Particularly, if $m_1 = 1$, $c_p = 0$ and $r_2 = \infty$, k is infinite resulting in zero swing at the virtual ground node. In general,

$$k = \frac{C}{c_p+\dfrac{c_1+c_p}{g_{m2}r_2}+\dfrac{c_p}{g_{m2}r_1}}, \qquad \text{if } \mathrm{m_1} = 1 \tag{5.13}$$

$$= \frac{C}{c_p+\dfrac{c_1+c_p}{g_{m2}r_2}+\dfrac{c_2+c_p}{g_{m2}r_1}}, \qquad \text{if } \alpha = \mathrm{m_1} = 1 \tag{5.14}$$

With this information, the nonlinear integrator can be modeled as before using a nonlinear function ($f(x) = x - \beta x^3$) preceding a linear integrator, where

$$f(x) = x - \frac{2g_{31}}{g_{m1}(2+kg_{m1}R)^3}x^3 - \frac{2g_{33}/(g_{m2}r_1)}{g_{m1}(2+kg_{m1}R)^3}x^3 \tag{5.15}$$

As discussed in Chapter 4 (Section 4.1), this can be modeled as an additional input ($2\beta V_{in}^3$) to an otherwise linear modulator. Thus, the third harmonic distortion at the modulator output is

$$HD_3 = \frac{\beta A_{in}^2}{2} \tag{5.16}$$

provided the initial transients are negligible. This is indeed the case, since the opamp is assisted in supplying the current steps demanded by the feedback DAC. Fig. 5.6 illustrates this by comparing the step response observed at v_d with and without opamp

assistance. Some initial transients can be seen in this case since the opamp is not assisted to drive the capacitor, c_2 (or $\alpha = m_1 = 1$). In such cases, the effect of any residual transients can be accounted by computing the time-averages (α_i), as discussed in Chapters 3 (Section 3.1) and 4 (Section 4.2).

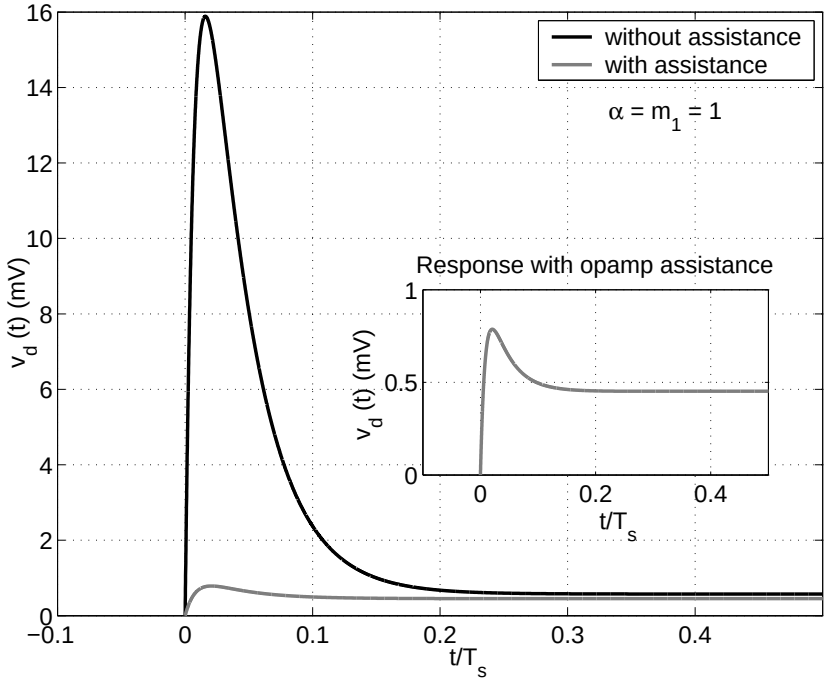


Figure 5.6: Step response of the active-RC integrator with and without opamp assistance ($\alpha = m_1 = 1$). Inset shows a zoomed in view of the response with assistance. g_{m1} = 27.12 μS, R = 150 kΩ, C = 5.77 pF, r_1 = 20 MΩ, c_p =10 fF, g_{m2} =100 μS, g_{m3} = 320 μS, c_1 =45 fF, r_2 = 1 MΩ.

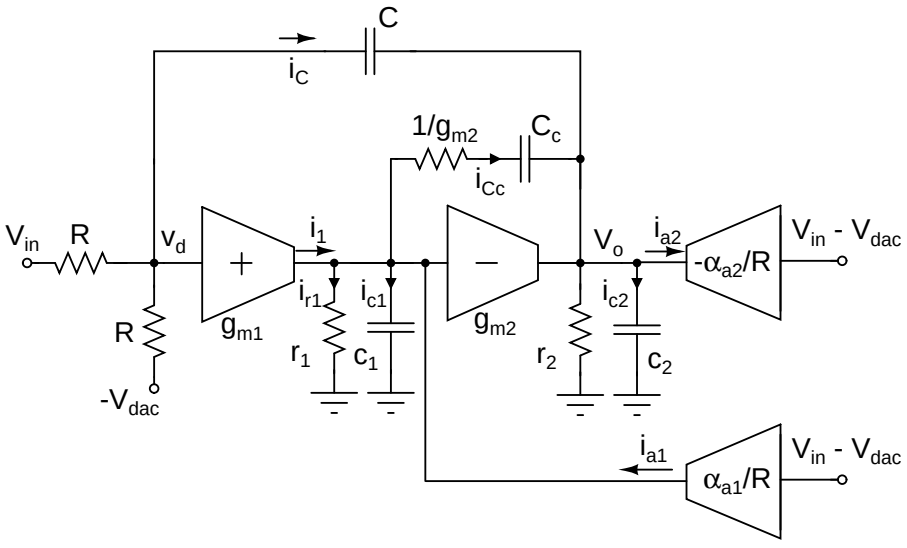


Figure 5.7: Assisted active-RC integrator with Miller compensated opamp

The corresponding model for a nonlinear Miller opamp based integrator can be obtained by a similar analysis, considering the circuit shown in Fig. 5.7. The assistants for each stage of the opamp are characterized by the parameters, α_{a1} and α_{a2} with

$$\alpha_{a1} = m_{11}\left(\frac{C_c}{C}\right) \tag{5.17}$$

$$\alpha_{a2} = m_{12}\left(1+\frac{C_c+c_2}{C}\right) \tag{5.18}$$

where m_{11} and m_{12} denote the relative variation of the respective assistant currents from their expected value. Considering the assistance provided to the first stage and assuming $i_C \approx (V_{in}-V_{dac})/R$, we have

$$i_1 = i_{Cc}+i_{c1}+i_{r1}-i_{a1} \tag{5.19}$$

$$= i_C\frac{C_c}{C}\left(1-m_{11}+\frac{1}{g_{m2}r_2}\right)+i_C\frac{c_1}{C}\left(\frac{1}{g_{m2}r_2}\right)+i_{r1} \tag{5.20}$$

With the second stage of the opamp also being assisted, the modified i_{r1} can be determined by proceeding along the same lines as with the feedforward compensated opamp. It can be shown that

$$i_1 \approx i_C\frac{C_c}{C}\left(1-m_{11}+\frac{1}{g_{m2}r_2}\right)+i_C\frac{c_1}{C}\left(\frac{1}{g_{m2}r_2}\right) + \frac{i_C(1-m_{12}+(C_c+c_2)(1-m_{12})/C)-i_3}{g_{m2}r_1} \tag{5.21}$$

Using the above expression, the nonlinear integrator can be modeled with a linear integrator preceded by the nonlinear function,

$$f(x) = x-\frac{2g_{31}}{g_{m1}(2+kg_{m1}R)^3}x^3 = x-\beta x^3 \tag{5.22}$$

$$\text{where,}\quad k = \frac{C}{C_c(1-m_{11})+\frac{c_1+C_c}{g_{m2}r_2}+\frac{(C+C_c+c_2)(1-m_{12})}{g_{m2}r_1}}$$

Recall that the feedforward opamp based integrators are more linear than their Miller counterparts due to the inherently large value of $k \approx C/c_p$. With opamp assistance, (in particular $m_{11} = 1$) it is evident from the above equation that large values of k (or smaller β) can be achieved with Miller opamp based integrators as well. In conjunction with reduced initial transients (due to $m_{12} = 1$), comparable performances can thus be achieved with a Miller opamp based integrator similar to that achieved with a feedforward opamp based integrator. By the same argument, it should be apparent that the linearity of a single-stage transconductor based active-RC integrator[4] can also be significantly improved with opamp assistance.

5.3 Nonidealities of the assistant blocks

The discussion so far explained how the linearity requirements of an opamp can be relaxed to enable the design of high performance single-bit CTDSMs. While assistance can help in reducing the power consumption in the opamp, two additional blocks have been introduced to accomplish this. These blocks were assumed to be ideal in the above discussion. The practical nonidealities of these assistant blocks (like noise and nonlinearity) can potentially degrade the modulator performance. This can result in significant power requirements in these assistants, which can make the overall power consumption of the assisted integrator more than that of the conventional integrator. However, as will be explained below, the nonidealities of the assistant blocks are not a serious concern. Hence, power can be conserved while designing these blocks.

[4] whose linearity can in principle be increased only through g_m or R

5.3.1 Effect of noise due to opamp assistants

The assistant circuitry can inject thermal and 1/f noise apart from supplying the desired signal current. Considering each of the integrator topologies shown in Fig. 5.4(a), (b) and (c), an intuitive understanding on the effect of these noise components can be obtained as follows. In all the cases, the spectral density of the noise current injected by the assistants is assumed[5] to be $8kT/R$.

In the single-stage transconductor case with assistance (Fig. 5.4(a)), the noise source of the assistants appears in parallel with the noise of the transconductor. In a well designed integrator, $G_mR \gg 1$. Therefore, it can be seen that the noise injected by the assistant can only slightly degrade the input referred opamp noise (which is much smaller than the input referred integrator noise).

With the two stage opamp using feedforward compensation (Fig. 5.4(b)), the noise current due to the assistants appears in parallel with that of the second stage of the opamp. Similar to the noise current of g_{m2}, the noise due to the assistant is negligible when referred to the opamp input.

For the Miller compensated two-stage opamp (Fig. 5.4(c)), the noise injected by the assistant circuitry appears in parallel with that injected by g_{m1} and g_{m2}. The noise injected at the opamp output is of little consequence, like in the feedforward example. The noise source occurring at the output of g_{m1} could be of potential concern. In a well designed integrator, where the gain-bandwidth product of the opamp (g_{m1}/C_c) is chosen to be much higher than the unity gain bandwidth of the integrator ($1/RC$),

$$g_{m1} \gg \frac{1}{R}\frac{C_c}{C} \tag{5.23}$$

[5]Though only thermal noise is considered, the discussion holds good for flicker noise as well.

The RHS of the above equation is the transconductance of the assistant block at the output of the first stage as shown in Fig. 5.4(c). This indicates that the noise penalty due to this assistant is also negligible.

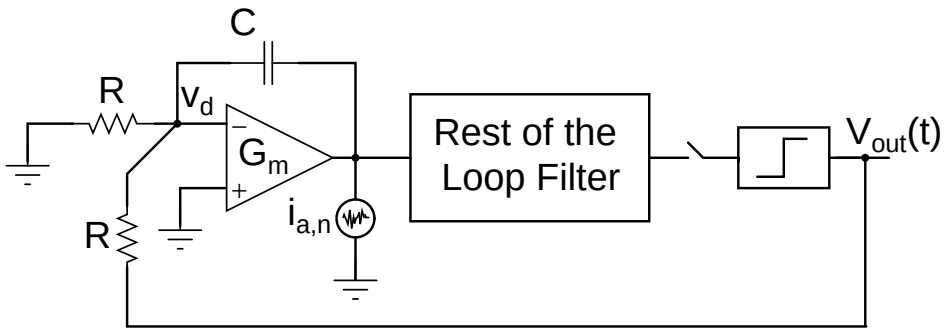


Figure 5.8: Simplified diagram of the modulator used to evaluate the effect of noise of the assistant transconductor

To obtain a quantitative understanding of the effect of the noise due to the assistants, consider the simplified representation of the CTDSM as shown in Fig. 5.8. Let $i_{a,n}$ denote the noise current injected by the assistant (with a spectral density $S_{ia,n}(f)$). Let G_m denote the low frequency transconductance of the opamp. Very little (≈ 0) current can flow through the integrating capacitor (C) at low frequencies[6]. If the rest of the modulator is assumed not to load the first integrator, the noise current of the assistant gets "absorbed" by the transconductor (G_m). This results in a corresponding swing in the virtual ground node given by[7]

$$v_d = \frac{i_{a,n}}{G_m} \tag{5.24}$$

Applying KCL at v_d (and noting that the current through $C \approx 0$), the output of the modulator is given as

$$V_{out} = \frac{2i_{a,n}}{G_m} \tag{5.25}$$

[6]assuming a high unity gain bandwidth for the first integrator
[7]assuming a deterministic $i_{a,n}$ for calculation purposes

Using this relation, the output noise spectral density can be related to $S_{ia,n}(f)$ as

$$S_{Vout,n}(f) = \frac{4}{G_m^2} S_{ia,n}(f) \tag{5.26}$$

The corresponding output noise spectral density of the modulator, assuming the input resistor's noise PSD to be $S_{iR,n}(f)$, is given by

$$S_{Vout,n}(f) = R^2 S_{iR,n}(f) \tag{5.27}$$

Since $S_{ia,n}(f)$ is of the same order as $S_{iR,n}(f)$, eq. 5.26 and 5.27 indicate that the noise of the assistant blocks is attenuated by a factor of $(G_m R/2)^2$ when compared to the noise contributed by the input resistor. Thus, by ensuring a large low frequency transconductance for the opamp, the noise of the assistant blocks can be made negligible at the output of the modulator.

5.3.2 Effect of distortion from the assistant circuitry

A similar analysis can show that the distortion introduced by the assistant circuitry is also attenuated by a large factor. If the assistant transconductor is assumed to be weakly nonlinear with a current function modeled as $i_a = V_{in}/R - g_{3a}V_{in}^3$ (where $V_{in}/R \gg g_{3a}V_{in}^3$), the amplitude of the third harmonic current tone injected by the assistant is

$$I_{3a} = \frac{g_{3a}A_{in}^3}{4} = HD_{3,a}\frac{A_{in}}{R} \tag{5.28}$$

where A_{in} denotes the input amplitude and $HD_{3,a}$ is the third harmonic distortion observed in the assistant current. This must result in a third harmonic component in the

virtual ground voltage v_d with an amplitude

$$A_{vd,3} = \frac{I_{3a}}{G_m} = \frac{g_{3a}A_{in}^3}{4G_m} \tag{5.29}$$

With an STF = 1, the signal amplitude at the modulator output is A_{in} and thus, the third harmonic distortion will be given as

$$HD_3 = \frac{g_{3a}A_{in}^2}{2G_m} = \frac{2HD_{3,a}}{G_m R} \tag{5.30}$$

It can be seen that the distortion of the assistant transconductor is also reduced by a factor $G_m R/2$. Since the assistant's distortion is not a serious concern, a class-AB architecture can be used to realize the assistant transconductor, leading to a power-efficient design. Also, it is easier to design the current steering DAC ($\pm V_{ref}/R$) for low distortion operation.

In a nutshell, it is seen that higher the low frequency transconductance of the opamp, better is the attenuation of the noise/nonlinear current components of the assistant blocks. Among the opamp topologies, the two-stage architectures can realize large low frequency transconductance[8] with the least power expenditure. Therefore, integrators with two-stage opamps will be more tolerant to the nonidealities of the assistants, when compared to their single-stage counterparts.

5.4 Simulation results

The models derived in Section 5.2 were confirmed by employing opamp assistance in the integrators of the third order single-bit CTDSM simulated in Chapter 4. Fig. 5.9

[8]The effective G_m is $g_{m1}r_1g_{m2}$ at low frequencies. However, in a Miller compensated opamp, the value of G_m reduces with a first order roll-off for frequencies above $1/\big(r_1(C_c + c_1)\big)$. This may not be critical as long as the signal bandwidth is smaller or $g_{m1}r_1g_{m2}$ is adequately large.

shows the variation in the harmonic distortion observed for the case of the feedforward compensated opamp[9] with different integrating capacitors. The harmonic distortion predicted using the models can be seen to closely match the simulated performance.

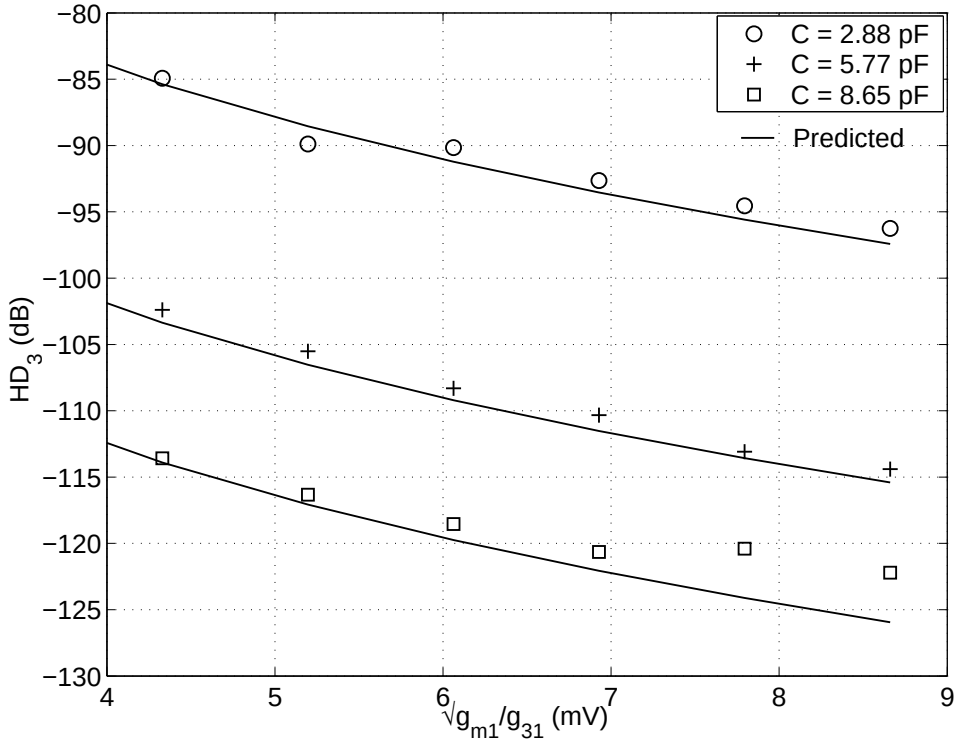


Figure 5.9: Variation of HD_3 with feedforward compensated opamp with assistance.

When compared to the modulator performance without assistance (shown in Fig. 4.10), the modulator with opamp assistance can be seen to tolerate stronger levels of nonlinearity (or higher values of $\sqrt{g_{m1}/g_{31}}$). More importantly, with negligible initial transients, the HD_3 is now inversely proportional to k^3 or C^3 as predicted by eq. 5.13. Note that a deviation is seen between the predicted and observed distortion, for the case of C = 8.65 pF. This is because the harmonic component becomes indistinguishable with the noise floor of the modulator at such low levels of nonlinearity. Fig. 5.10 shows the corresponding variation in HD_3 observed with the Miller compensated opamp, when assistance is provided only to the second stage ($m_{11} = 0$, $\alpha_{a2} = m_{12} = 1$).

[9]In all the simulations, $\alpha = m_1 = 1$ since $C \gg c_2$.

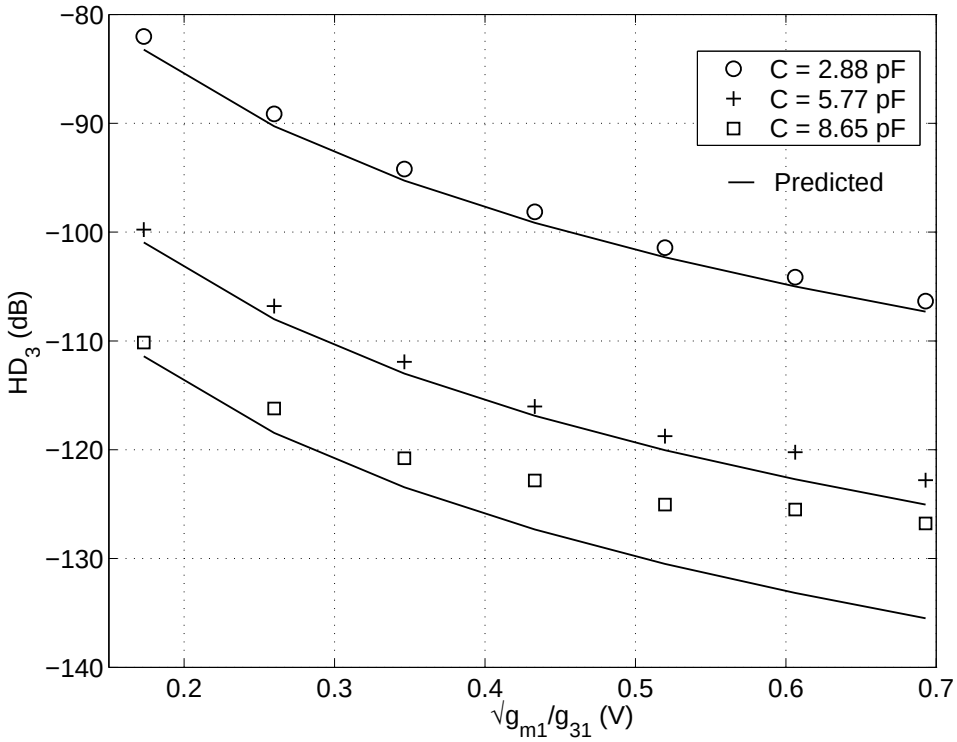


Figure 5.10: Variation of HD_3 with Miller compensated opamp with assistance provided only to the second stage ($m_{11} = 0$, $\alpha_{a2} = m_{12} = 1$).

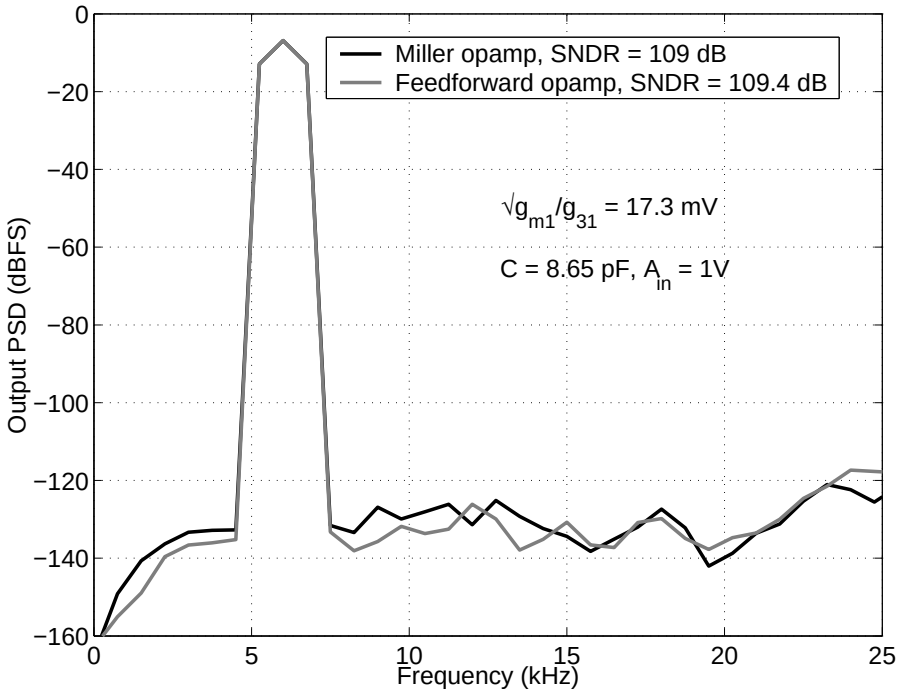


Figure 5.11: In-band PSD of the single-bit CTDSM employing a nonlinear active-RC integrator with both Miller and feedforward compensated opamps. Assistance has been provided to both the stages of the Miller opamp.($m_{11} = 1$, $\alpha_{a2} = m_{12} = 1$).

Referring to the modulator performance without assistance (Fig. 4.9), a slight improvement can be seen in the performance with such an assistance. This can be attributed to the reduction in the initial transients. However, if the opamp's input transconductor is also assisted ($m_{11} = 1$), performances comparable to that of the feedforward opamp can be expected. Fig. 5.11 illustrates such an improvement, obtained by assisting both the stages of the Miller compensated opamp. For comparison, the in-band PSD observed with the feedforward opamp based integrator is also provided.

5.5 Application of opamp assistance to multi-bit modulators

The concept of opamp assistance can also be applied to reduce the effect of integrator nonlinearity in multi-bit modulators. In such cases, the linearity of the first integrator can be improved by choosing

- two-stage feedforward compensated opamps to significantly reduce β and
- opamp assistance to reduce the initial transients ($w(t)$)

While the opamp can be assisted in the same manner as discussed for the single-bit modulator (Fig. 5.2), a modified scheme can also be suggested. This can be understood from the fact that the initial transients are caused by the steps in the feedback DAC voltage given by $V_{dac}[n] - V_{dac}[n-1]$. In the technique discussed so far, the DAC current is entirely provided by the assistant to remove the need for the opamp to source/sink steps of DAC current. Alternately, the high speed current requirements on the opamp can be relaxed by supplying only the "incremental" steps of the DAC current, $(V_{dac}[n] - V_{dac}[n-1])/R$ in each clock cycle. Illustrated in Fig. 5.12, such a scheme reduces the number of unit DAC cells in the assistant DAC, since the transitions typically

span only 2 or 3 LSBs. This can be advantageous in a multi-bit design in terms of the area and power consumed by the assistant circuitry. However, with this modified assistance, the opamp has to sink the current difference, $i_C - i_a = (V_{in}[n] + V_{dac}[n-1])/R$. This can be modeled as an error current injected by the assistant DAC. Though it increases the opamp current, this does not result in performance degradation, as the errors get absorbed by the opamp. It is sufficient for the current in the opamp to be adequate enough to sink these currents.

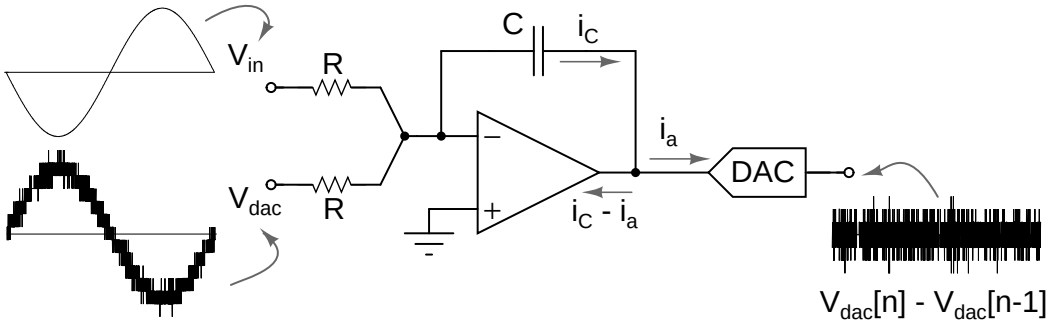


Figure 5.12: Modified assisted opamp integrator in a multi-bit CTDSM

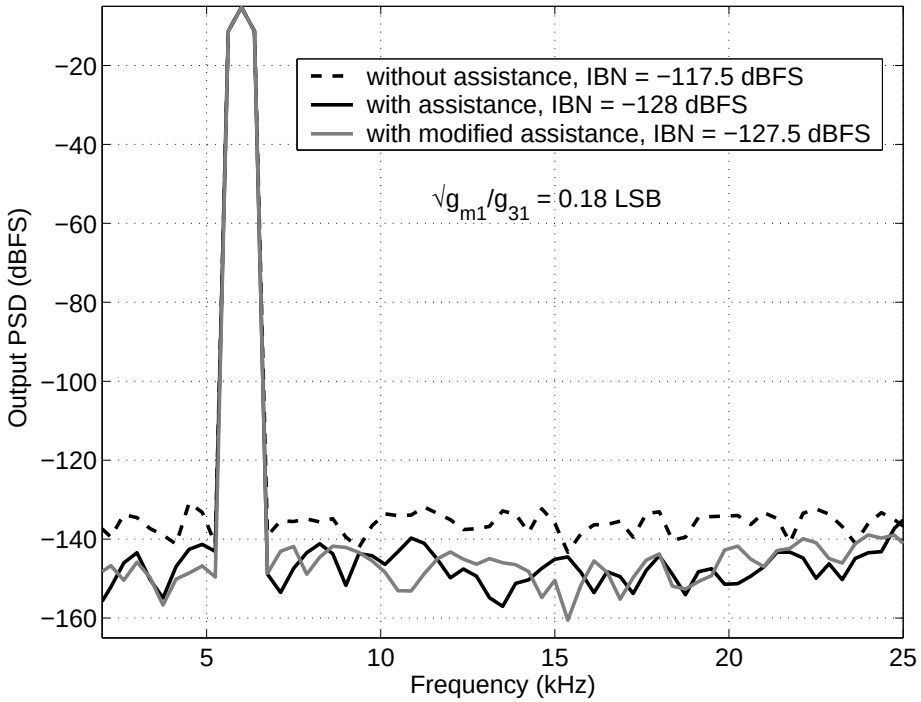


Figure 5.13: Comparison of the in-band PSD of the output of a multi-bit CTDSM employing a nonlinear active-RC integrator having a feedforward compensated opamp, with and without opamp assistance. The PSD observed with modified opamp assistance is also shown.

Fig. 5.13 illustrates the improvement observed in the 4-bit CTDSM (simulated in Chapter 2) with a nonlinear feedforward opamp based active-RC integrator. It is clearly seen that both the proposed schemes of opamp assistance aid in improving the modulator performance.

5.6 Prospects and related work

The assisted opamp technique described in this chapter is fundamentally a way of improving the performance of an integrator. It should benefit any modulator topology (like CIFF or CIFB) if the critical integrators are assisted with this technique. While the idea has been explained for an NRZ DAC, the benefits carry over to other DAC pulse shapes as well (like RZ, SCR DACs). This is because the cancellation of the swing at the virtual ground of the opamp is independent of the integrator input waveforms (as long as there is good matching between the assistant and the net input current flowing into the integrating capacitor).

The idea of "helping" an opamp has been attempted before, in the context of a switched capacitor integrator (Stevens and Miller (1994)). In that work, the authors use a booster circuit to supply a large portion of the (initial) current required in an SC integrator. This is done by using the knowledge of the "sampled" input of the integrator. By this way, the slew-rate of the integrator is shown to be enhanced. While this might be suitable for a switched capacitor application (since only the settled value is of interest), it can be problematic for a continuous-time design.

CHAPTER 6

DESIGN OF LOW POWER SINGLE-BIT AUDIO MODULATORS

The recent delta-sigma modulators targeting the audio range have been built using continuous-time loop filter (Gerfers *et al.* (2003) Ortmanns *et al.* (2005) Baggini *et al.* (2006) Nguyen *et al.* (2005) Dorrer *et al.* (2006) Pavan *et al.* (2008)). In all these works, power reduction has been the main motivation for choosing CTDSMs over their discrete-time counterparts. While the first three designs (cited above) have used a single-bit quantizer, subsequent designs have largely employed multi-bit quantizers. This is because multi-bit designs are less sensitive to clock jitter and relax the linearity requirements of the loop filter, when compared to their single-bit counterparts. Also, the inherent quantization noise of multi-bit quantizers is smaller and with aggressive NTFs being possible, the in-band quantization noise gets further reduced.

Though single-bit modulators are associated with such design issues, they do offer some attractive features.

- A single-bit quantizer greatly reduces the power consumed by the digital section of the modulator - like the comparators, digital logic and the clock drivers.
- A single-bit quantizer reduces the design area, owing to the smaller number of comparators/DAC elements.
- The fact that the Dynamic Element Matching (DEM) circuitry is not needed, further reduces area and complexity.
- Though not a serious concern in audio modulators, the absence of the DEM block can significantly reduce the excess loop delay.

To illustrate the above mentioned advantages with an example, in the multi-bit design reported by Pavan *et al.* (2008), about 30% of the modulator power is consumed

by the digital section - mainly by the clock generator circuitry that has to drive the 4-bit internal flash ADC and the DEM block. In addition, one-fourth of the modulator area is found to be occupied by the Flash, DEM and DAC circuitry. This clearly signifies the power and area advantage possible when a single-bit design is adopted. This was the motivation behind designing single-bit audio modulators, the design of which will be discussed in this chapter. The modulators were targeted for the same 15-bit performance as the above mentioned multi-bit design, with a similar power dissipation.

6.1 Addressing the issues of a single-bit design

The issues associated with a single-bit design should first be addressed in order to leverage its implementation advantages. The problems have been circumvented by adopting the following system and circuit level choices.

A.) In-band quantization noise : The increase in quantization noise with the use of a single-bit quantizer is countered by an equivalent increase in the Over-Sampling Ratio[1]. Thus, an OSR of 128 has been chosen, resulting in a sampling frequency of 6.144 MHz for 24 kHz signal bandwidth. Complex zeroes are spread in the signal band to further minimize the in-band quantization noise power. With a third order maximally flat NTF having an OBG of 1.5, this results in an in-band peak SQNR of about 110 dB . This is well above the desired 92 dB SNR of the modulator, which is ensured to be determined purely by device (thermal) noise. In addition, the dynamic range of the modulator is maximized by using a full scale voltage of 3.6 V (peak-to-peak differential) for the quantizer, which is the maximum possible with the given supply voltage.

[1]The increase in the modulator's power consumption (due to the increase in the digital power consumption) is only marginal with an increase in OSR. This is because the loop filter is the major power dissipating block in a single-bit CTDSM.

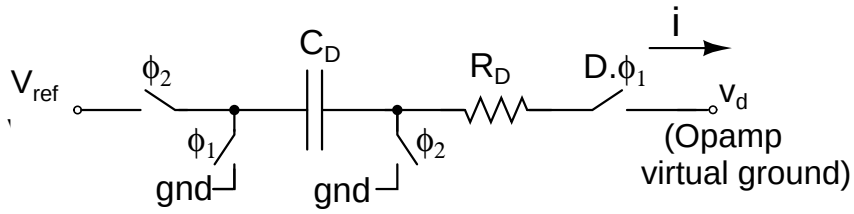
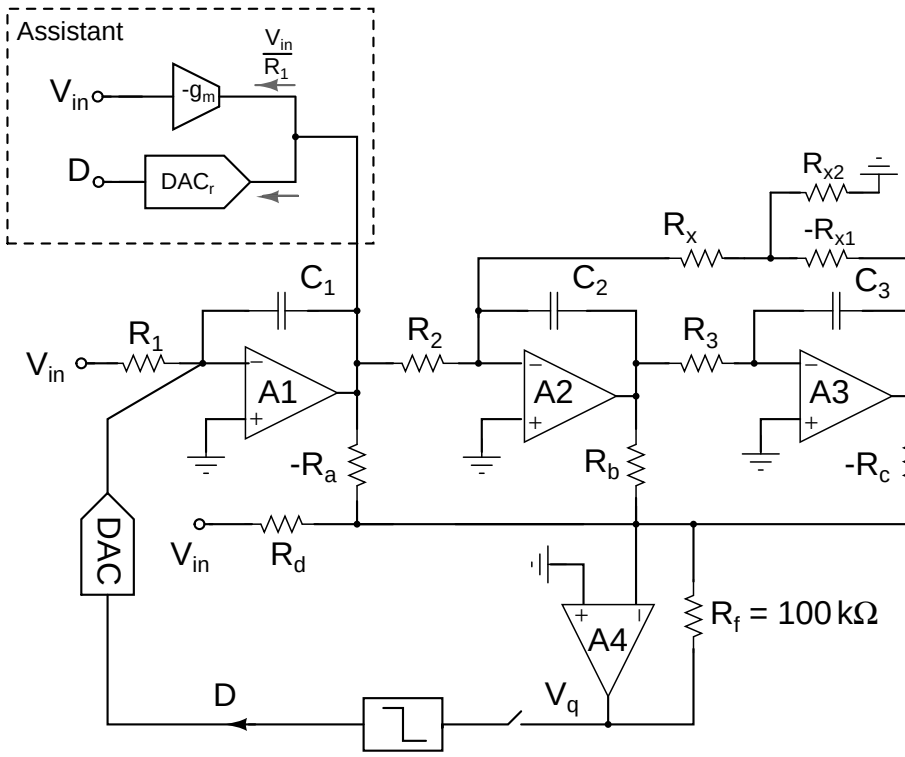


Figure 6.1: Single-ended schematic of an SCR DAC

B.) Clock jitter : It is well known that jitter affects the performance of CTDSMs by perturbing the width of the fedback DAC pulse, making an NRZ DAC preferable to an RZ DAC. With an NRZ DAC, this issue can be addressed through appropriate PLL design. Alternatively, an SCR DAC implementation (Ortmanns *et al.* (2005)) as shown in Fig. 6.1 can also be employed.

C.) Power consumption in the loop filter : The discussions in Chapter 3 clearly brought out the increased linearity requirements of the loop filter in a single-bit CTDSM. The power required to design linear integrators can potentially offset the power advantage gained in the digital section. Further, in case of an SCR DAC, reduced jitter sensitivity is achieved only at the expense of increased linearity requirements of the first integrator. This can easily undermine the objective of designing a low power single-bit CTDSM. To address this problem, the technique of opamp assistance is used. With the help of this technique, as targeted, 15-bit performance has been in demonstrated in single-bit CTDSMs based on NRZ and SCR feedback DACs, with a power consumption that is comparable to that of the multi-bit modulator reported by Pavan *et al.* (2008). Implemented in 0.18 μm CMOS technology, the NRZ/SCR modulators consume 110/122 μW from a 1.8 V supply and achieve dynamic ranges of 92.5/91.5 dB for a signal bandwidth of 24 kHz. The architectural and circuit details of these modulators form the subject of the rest of the chapter.

6.2 Loop filter architecture



NRZ DAC Design		
$R_a = 235\,k\Omega$	$R_b = 325\,k\Omega$	$R_c = 490\,k\Omega$
$C_1 = 5.43\,pF$	$C_2 = 550\,fF$	$C_3 = 1.02\,pF$
$R_{x2} = 7.9\,k\Omega$	$R_d = 1.0\,M\Omega$	

SCR DAC Design		
$R_a = 232\,k\Omega$	$R_b = 400\,k\Omega$	$R_c = 915\,k\Omega$
$C_1 = 6.67\,pF$	$C_2 = 400fF$	$C_3 = 600\,fF$
$R_{x2} = 4.1\,k\Omega$	$R_d = 1.6\,M\Omega$	

Figure 6.2: Modulator architecture and component values (DAC is either NRZ or SCR)

A single-ended equivalent of the modulators showcasing the loop filter architecture along with the respective component values[2] is shown in Fig. 6.2. The loop filter is implemented as a CIFF structure. For low noise and high linearity, active-RC techniques are used to implement the integrators. Opamp assistance is used only in the first integrator, since distortion introduced by subsequent circuitry gets attenuated by the high in-band gain of the first integrator. This is an inherent advantage of the CIFF architecture, and the high OSR (= 128) being used. For simplicity, assistance is not provided for the first opamp to drive the capacitive loads (≈ 300 fF), due to the relatively larger value

[2]The inversion of the signals have been indicated using the negative sign used with the resistors.

for the first integrating capacitor (C_1). The first integrator uses 150 $\mathrm{k}\Omega$ resistors (to limit thermal noise), while 600 $\mathrm{k}\Omega$ resistors are used in the second and third integrators, to reduce power dissipation. Some important aspects of the architecture can be elaborated as follows.

a.) Direct path to the quantizer : A direct path is added from the modulator input to the loop filter output through the resistor R_d. This technique, originally proposed to reduce distortion in discrete-time DSMs (Silva *et al.* (2001)), has been adopted in this design for the following reason. In order to limit the swing at the third integrator's output, large values for the third integrating capacitor (C_3) are required. With the direct path, the signal component of the loop filter output is provided by the modulator input. Therefore, relatively smaller values can now be used for the third integrating capacitor. It needs to be mentioned that the drawback of having a direct path is the reduced anti-aliasing property of the CTDSM. However, as explained in Appendix H, the degradation in the anti-alias rejection is tolerable for practical purposes.

b.) Implementation of the notch in the NTF : Realizing NTF complex zeros (which improve the SQNR by about 8 dB) is traditionally done using a single feedback resistor around two integrators (in the notation of Fig. 6.2, $R_{x1} = 0$ and $R_{x2} = \infty$). For a third order NTF and 24 kHz bandwidth, calculations show that using a single resistor to realize the notch necessitates an impractically large R_x (of approximately 200 MΩ!). To avoid this, the notch is implemented using a T-network formed by $R_x = 2\,M\Omega$, $R_{x1} = 1\,M\Omega$ and R_{x2}. The notch frequency is given by $\omega_z^2 \approx (R_{x2}/R_{x1})\,(1/R_3 R_x C_2 C_3)$. The resistors, R_{x1} and R_{x2} are realized without any power or area overhead, by re-using the common-mode detecting resistors employed at the output of the third opamp (A3).

c.) Realizing the summing operation in the loop filter : A CIFF loop filter has several advantages with respect to suppressing noise and distortion from sources other than the first integrator. However, feed-forward architecture requires a summing operation to be performed. This can be accomplished by either having capacitive feed-ins into the third integrator or by using a dedicated summing amplifier.

In the capacitive summation approach, the first integrator needs to drive a capacitive load. With the capacitive load current being proportional to that flowing through the integrating capacitor (C_1), this can potentially increase the power dissipation in the first opamp or the assistants, to ensure low distortion operation of the modulator. Further, the third integrator is now in the high speed path (the $1/s$ path of the loop filter) of the modulator. The feed-in capacitors needed to implement summation result in increased delay when compared to the delay of a stand alone integrator. This requires higher power to be burnt in the third opamp. Hence, even though capacitive summation needs only three opamps, it is not clear that it saves power. This was the reasoning behind using a separate summing amplifier (opamp A4).

d.) Use of resistor and capacitor banks : As discussed in Chapter 1, the variation in RC time constants of the loop filter (due to process shifts and changes in ambient temperature) can be problematic in a CTDSM design. To counter this, the integrating resistors and capacitors are implemented as digitally tuned banks as shown in Fig. 6.3. While the time constant variations can be compensated by tuning 'only' the capacitors or the resistors, the resistor and capacitor banks are tuned individually, in this design. This alleviates the need to increase the value of the resistances (in particular R_1) when the capacitors are small, which prevents any increase in the thermal noise. The switches S_{0-7} are externally controlled and can compensate for RC shifts of up to $\pm 25\%$ from the

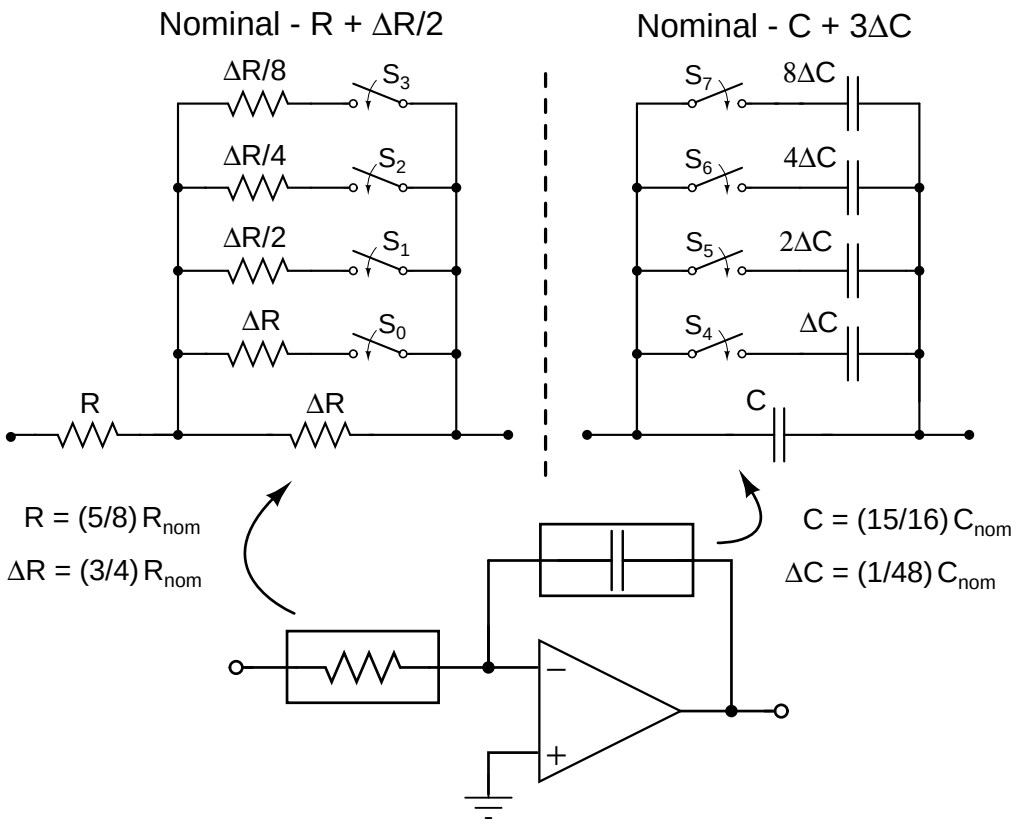


Figure 6.3: Resistor and capacitor banks used in the integrators

nominal value and bring them to within $\pm 5\%$. Under nominal conditions, the desired resistance and capacitance are realized by turning on switches S_0 and S_{4-5}.

e.) Scaling of the NTF : Finally, thanks to the single-bit quantizer, the output of the summing amplifier is scaled. In essence, the output of the loop filter is scaled by a factor of 8 in both the modulators without affecting the NTF. This has been used to advantage in two ways - the power consumption of the loop filter is minimized and also, a high speed opamp architecture which cannot support full rail swings, is used to design the opamp A4. For the same reason, a large value is chosen for the first integrating capacitor (C_1) to limit the power dissipation and output swings of the first two integrators (realized by opamps, A1 and A2).

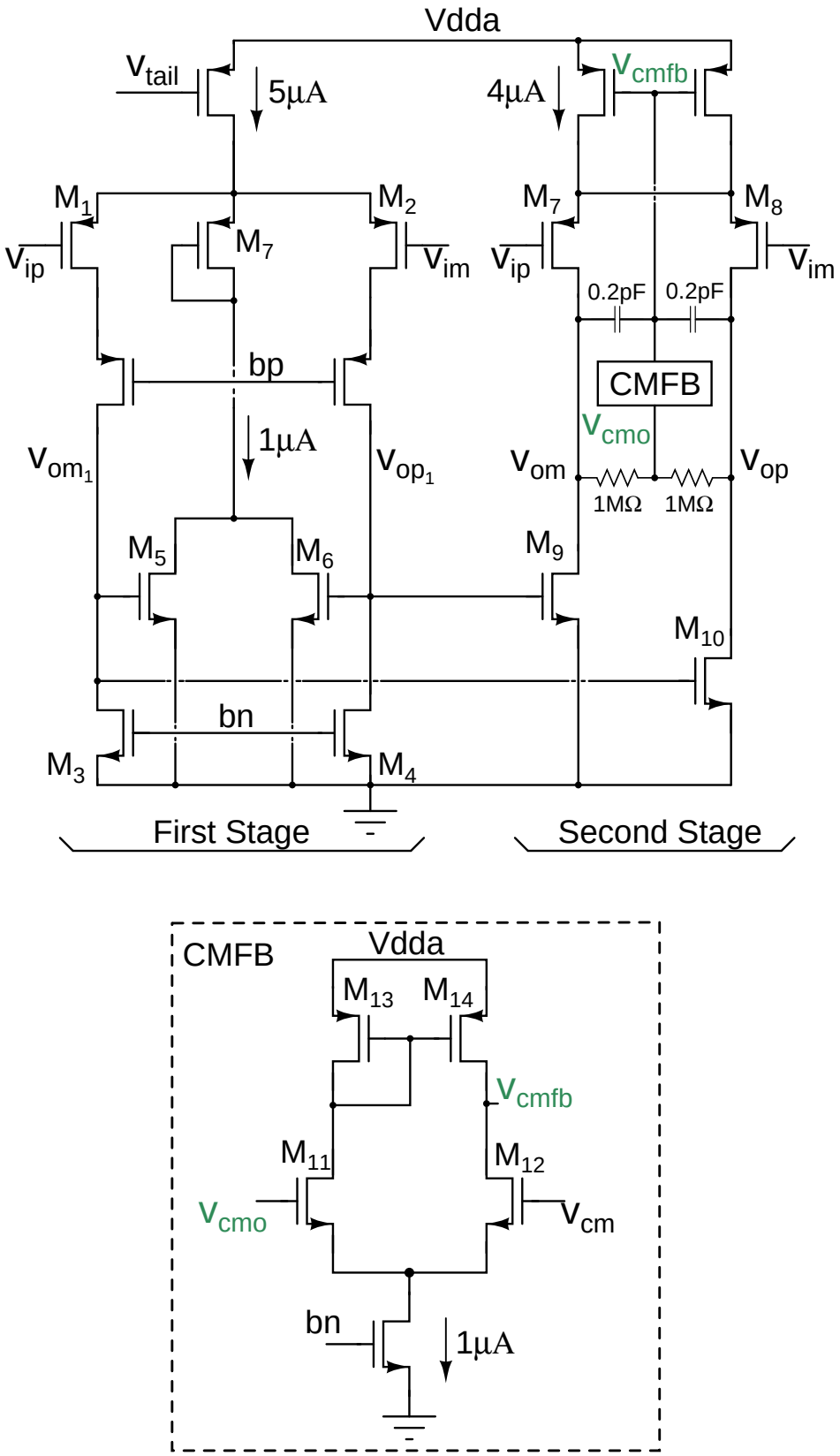


Figure 6.4: Operational amplifier used in the first integrator

6.3 Circuit design details

6.3.1 Operational amplifiers

The opamps in the active-RC integrators are realized using two-stage architectures with feedforward compensation, to achieve better speed and linearity. The opamp used in the first integrator of the NRZ DAC based modulator is shown in Fig. 6.4. The first stage consists of long channel devices M_1, M_2, M_3 and M_4 to lower the input referred 1/f noise. M_5 and M_6 provide common-mode feedback for the first stage. These are minimum length devices, carrying 20% of the first stage tail current. Since $\mathrm{M}_{9,10}$ and $\mathrm{M}_{5,6}$ have the same V_{DS}, the quiescent current in the second stage is accurately set to $4\,\mu$A by choosing M_9 to be 8 times wider than M_5. The reason for the choice of $4\,\mu$A is the following. The modulator coefficients are scaled in such a manner that $1.5\,\mu$A is needed from the first opamp to drive the resistive load current of the subsequent stages. $2.5\,\mu$A has been allocated for design margin. The total current drawn by the opamp is $15\,\mu$A. The simulated DC gain and unity gain bandwidth of the opamp are 65 dB and 55 MHz respectively.

A notable feature of the opamp is that the second stage current is re-used to realize the feedforward transconductance. This results in a power efficient opamp architecture, though with the drawback of increased output swing constraints. However, as discussed in the previous section, the first integrator is node scaled to ensure the swings are well-within the limits set by the given opamp topology.

A similar opamp topology can be found in Ouzounov *et al.* (2007), where resistive common-mode feedback has been used at the output of the first stage. In contrast, transistors are used to detect the common-mode voltage in this design. This occupies

a smaller area and more importantly, maintains the inherent high DC gain of the first stage. Further, the common-mode feedback scheme used allows to set the current density of $\mathrm{M}_{3,4}$, independent of that in the output stage. $\mathrm{M}_{3,4}$ are thus sized based on 1/f noise considerations, while $\mathrm{M}_{5,6}$ and $\mathrm{M}_{9,10}$ are optimized for speed.

The opamp used in the SCR DAC based modulator is similar to that in Fig. 6.4 - the only difference is that the quiescent current in the second stage is increased to $6\,\mu$A per leg. The reason for this is the mismatch expected between the assistant and feedback DAC currents, which will become clearer in the next section.

The opamps used in the subsequent integrators and the summing block are scaled down versions of the first opamp, since their noise and nonlinearity are of little concern. The current consumed by each of these opamps is $8.5\,\mu$A. Since 1/f noise is not an issue in these opamps, NMOS input and feedforward pairs are used, while g_{m2} is implemented using PMOS devices. As mentioned in the previous section, the second stage common-mode feedback resistors of the third opamp are modified to implement the T-network as shown in Fig. 6.5.

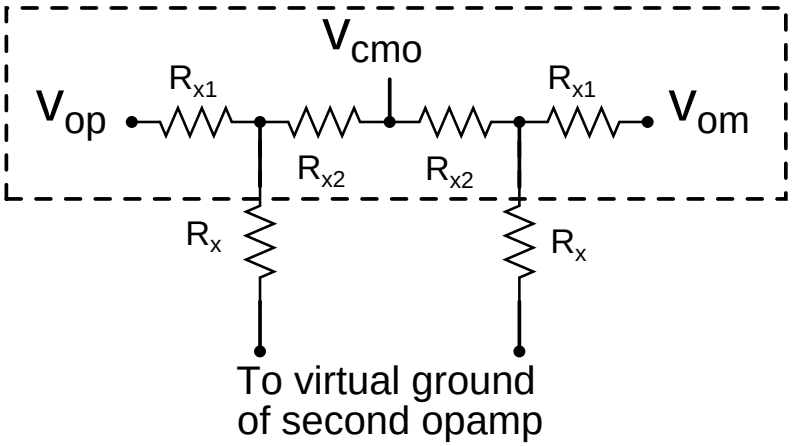


Figure 6.5: Common mode feedback sensing used in third opamp (R_{x1} and R_{x2}) to realize the T-network. Resistor R_x is also shown for convenience.

6.3.2 Quantizer

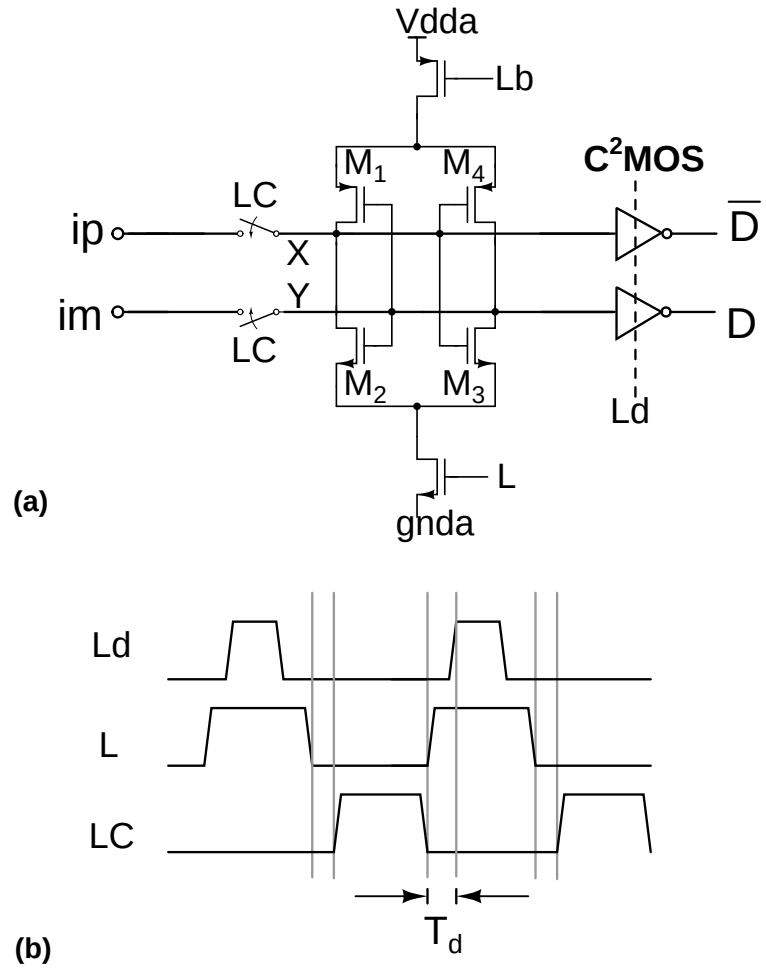


Figure 6.6: (a) Comparator schematic and (b) Clock waveforms

The circuit diagram of the comparator and the corresponding clock waveforms are shown in Figs. 6.6(a) and (b) respectively. The comparator operates as follows. When the latch-connect signal LC is high, the back-to-back inverters formed by M_1, M_2, M_3 and M_4 are disabled and the differential input is sampled on nodes X and Y. The regeneration phase begins thereafter, when the latch signal (L) goes high. The decision of the latch is held by two $\mathrm{C2MOS}$ (clocked CMOS) inverters, which sample the latch output after a small delay (T_d). The excess delay in the quantizer realized is about 100 ps. Since the NTF has a relatively low OBG of 1.5, this delay is of little consequence.

6.3.3 Feedback DACs

The NRZ DAC based design uses a simple resistive DAC ($R_1 = 150\,\mathrm{k\Omega}$) for low excess noise[3]. The schematic of the first integrator of the SCR DAC based modulator is shown in Fig. 6.7(a) and the details of the DAC are shown in Fig. 6.7(b).

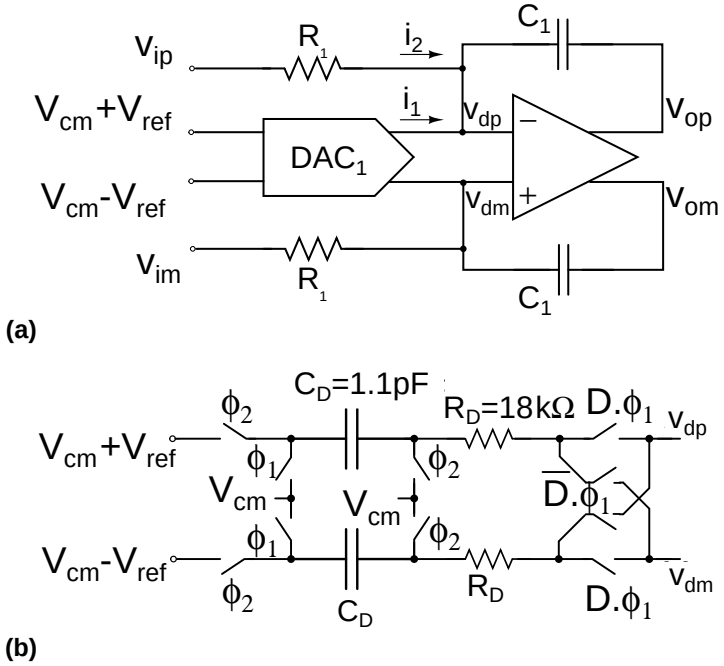


Figure 6.7: (a) Schematic of the first integrator in the modulator with the SCR DAC (b) Schematic of the differential SCR DAC

The capacitors C_D are charged to $\pm V_{ref}$ in phase ϕ_2. In ϕ_1, they are discharged into the virtual grounds (v_{dp} & v_{dm}) depending on the quantizer output, D. The discharge time constant is given by $(R_D + 2R_s)C_D$, where R_s denotes the resistance of the switches in series with C_D. As discussed in Chapter 1 (Sections 1.4.3 and 1.4.4), the choice of R_D involves a trade-off between linearity and jitter noise. In this design, $C_D = 1.1\,\mathrm{pF}$ and $R_D + 2R_s = 22.5\,\mathrm{k\Omega}$ are used, resulting in a time constant of about 25 ns. The exponentially shaped DAC current thus has a peak of $40\,\mu\mathrm{A}$, decaying to about $1.5\,\mu\mathrm{A}$ in half clock cycle (which is ≈ 3.3 time constants of the SCR DAC). The average current is $6\,\mu\mathrm{A}$, as with the NRZ feedback DAC.

[3] A comparison of resistive and current-steering DAC architectures is provided in Appendix E.

6.3.4 Assistant circuitry

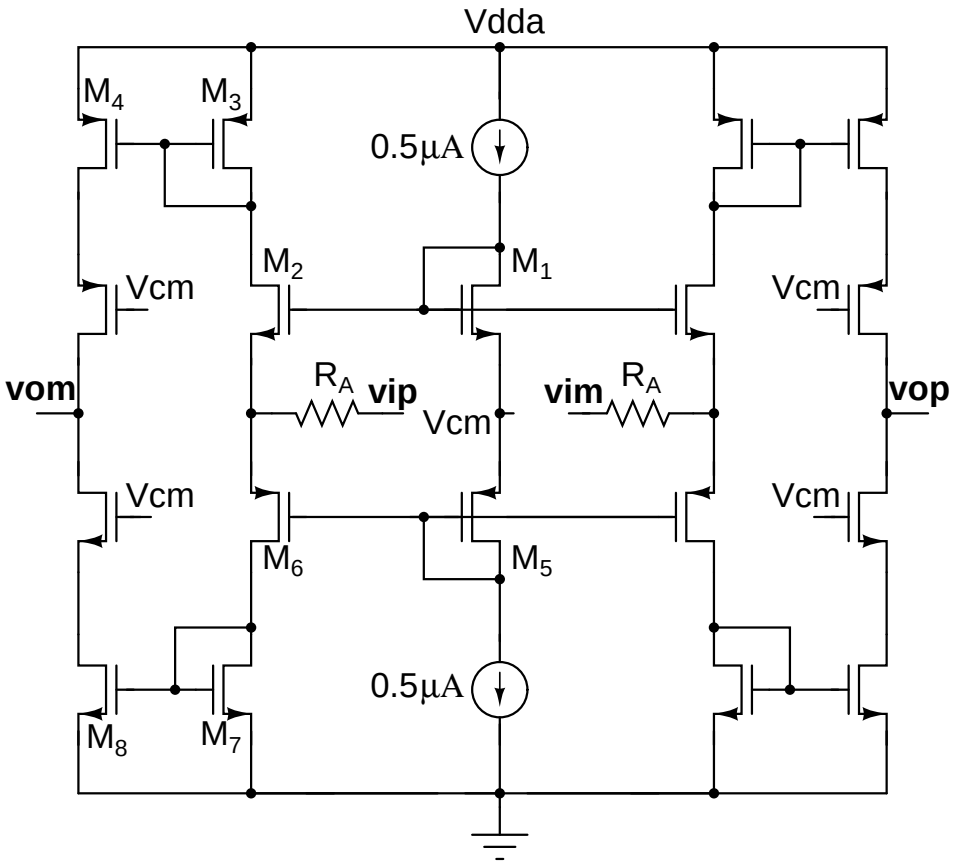


Figure 6.8: Transconductor used in the assistant circuit

Assistant transconductor : Fig. 6.8 shows the schematic of the assistant transconductor used in the first integrator. It is a class-AB design, comprising of complementary common gate stages M_2 and M_6 which carry 0.5 μA of current under no signal conditions. R_A is chosen to have a value of 300 kΩ. M_4 and M_8 are accordingly chosen to be twice as large as M_3 and M_7, to realize an effective small-signal transconductance of (1/150 k)S. As explained in Chapter 5 (Section 5.3.2), distortion components generated by this transconductor are absorbed by the first opamp. In this design, about 0.5$\%$ distortion from the assistant circuitry is found to be tolerable, which is easily achieved by this class-AB design.

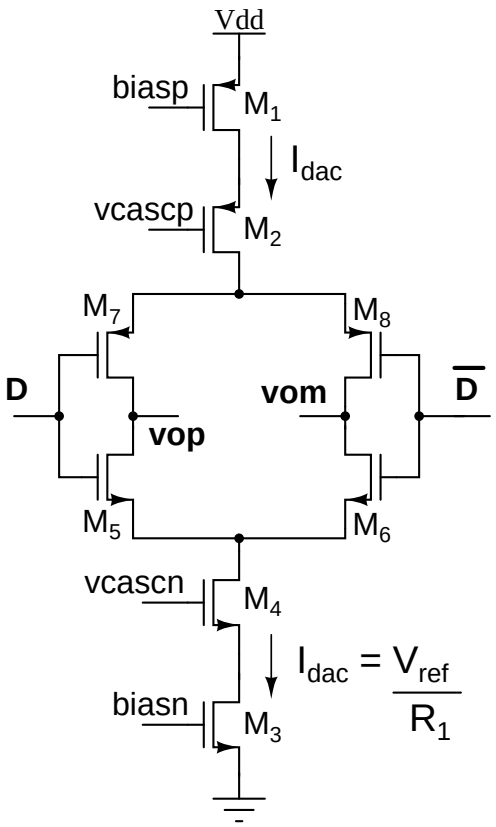


Figure 6.9: Assistant DAC used in the NRZ modulator

Assistant DAC in the NRZ design : The assistant DAC in the NRZ modulator consists of 6 μA cascoded NMOS and PMOS current sources, which are steered using differential pairs of switches (as shown in Fig. 6.9). The current sources are derived from a current proportional to V_{ref}/R_1, where R_1 (= 150 kΩ) denotes the resistance of the modulator feedback DAC. In this way, the magnitude of the assistant DAC current is ensured to track variations of the feedback DAC current, over process and temperature.

Assistant DAC in the SCR design : The assistant DAC of the SCR modulator needs to provide an exponentially decaying current with a peak of about 40 μA and a time constant as same as the feedback SCR DAC. More importantly, the assistant DAC should robustly track the variations in the feedback DAC current, over process and temperature. This represents a significant circuit design challenge. A possible choice for the

assistant DAC is the circuit proposed by the authors of Ortmanns *et al.* (2003). Their circuit is as shown in Fig. 6.10 and it operates as follows.

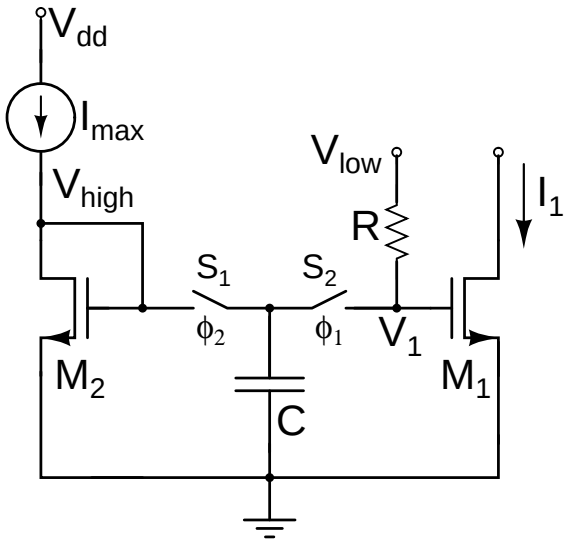


Figure 6.10: Simplified schematic of an exponentially decaying pulse generator proposed in Ortmanns *et al.* (2003)

For simplicity, M_1 and M_2 can be assumed to be identical devices. During phase ϕ_2, the capacitor C is charged to V_{high}. In ϕ_1, it is discharged to $V_{low} = V_{TN}$ (the threshold voltage of $\mathrm{M}_1/\mathrm{M}_2$) through a resistor R. The choice of V_{TN} for V_{low} is to avoid having to charge C all the way from zero during ϕ_2. If the switches are ideal, the gate voltage of M_1 (denoted by V_1) is an exponentially decaying pulse with a time constant $\tau = RC$. The drain current of M_1, which is initially I_{max}, decays to zero as V_1 approaches V_{TN}. If the transistors are assumed to behave like square-law devices, it can be seen that the time constant of the decaying current is 0.5τ. Several problems were encountered while attempting to adopt this technique in this design, as described below.

- *Error in the initial current :* When ϕ_1 goes high, V_1 should initially be V_{high} so that $I_1 = I_{max}$. Several mechanisms prevent this from happening. Charge sharing between C and the gate capacitance of M_1 causes a reduction in V_1, which is aggravated by the voltage drop across the switch, S_2. Threshold mismatch between M_1 and M_2 leads to an error proportional to the g_m of M_1 (since the initial current is high, the g_m and current error are also large).

- *Error in the decay time constant :* Since M_1 is not a true square-law device, the decay time constant of I_1 does not track the time constant of the feedback DAC pulse.
- *Error in the off state :* If V_{low} is larger than V_{TN} (due to mismatch), I_1 does not go to zero.

Excessive swing at the virtual ground of the opamp can be prevented only if there is good matching between the currents of the SCR feedback DAC and that of the assistant DAC. Understandably, a high degree of matching is required in the initial periods, when the current is at its peak. While such issues of the circuit of Fig. 6.10 are tolerable in the design of Ortmanns *et al.* (2003) (as is apparent from that reference), it results in significant degradation in the modulator performance when used as the assistant DAC.

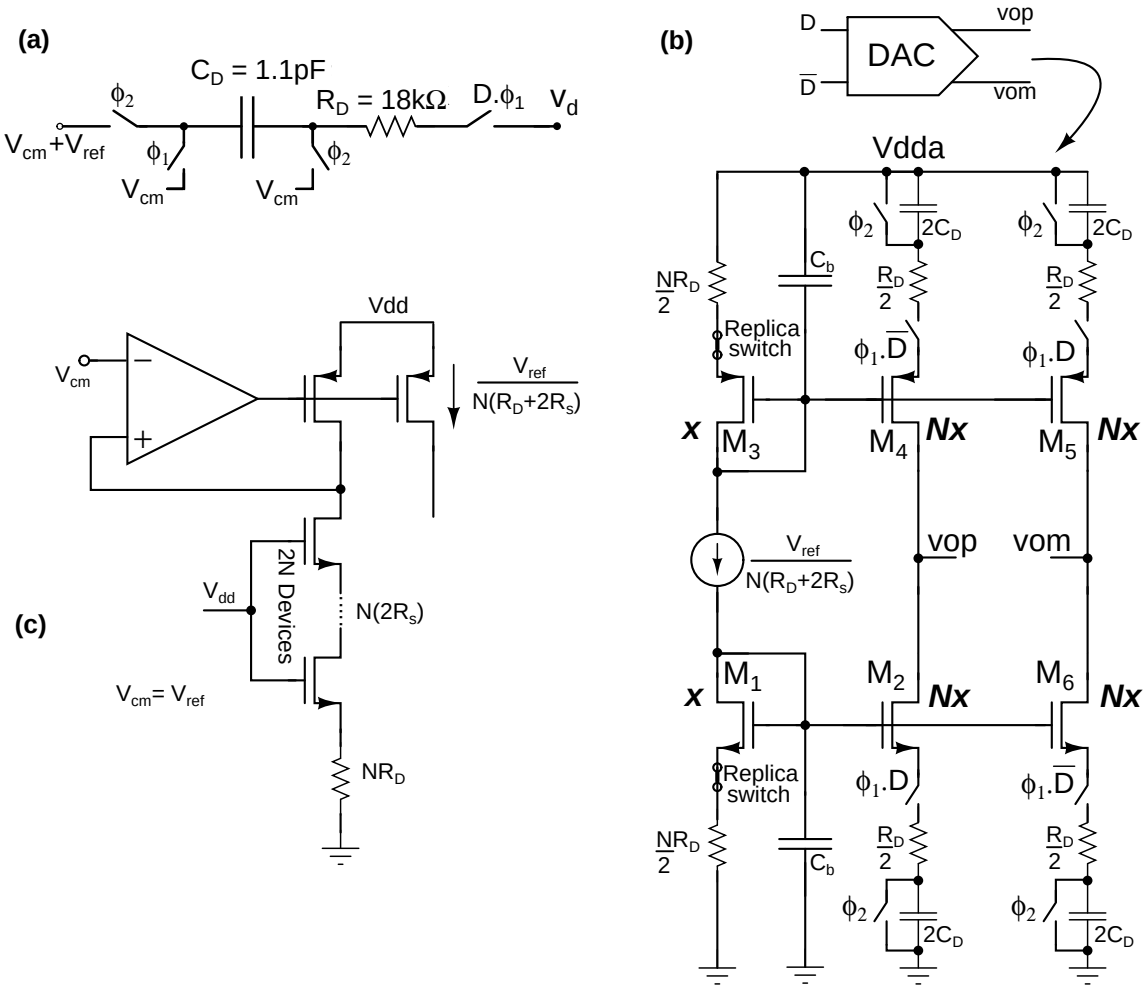


Figure 6.11: (a) Single branch of the SCR feedback DAC (b) Schematic of the SCR assistant DAC and (c) Simplified schematic of the bias current generator.

The fundamental reason for the inaccuracy of the circuit in Fig. 6.10 is that all error mechanisms directly impact the V_{GS} of M_1. This problem has been overcome in this modulator design by desensitizing the exponentially decaying current from the assistant DAC with respect to the properties of the transistor. This is achieved by switching the capacitor at the source of the device (akin to the benefits of source degeneration), rather than at its gate. The functioning of the proposed assistant circuit is explained with the help of Fig. 6.11. For reference, one branch of the SCR feedback DAC is shown in Fig. 6.11(a), while Fig. 6.11(b) shows the schematic of the assistant DAC.

The operation of the NMOS portion is described below - the PMOS section operates in a similar fashion. $\mathrm{M}_2, \mathrm{M}_6$ are N (=40) times wider than M_1, which is biased with $I_{ref} = V_{ref}/\big(N(R_D + 2R_s)\big)$ (generated by the circuit shown in Fig. 6.11(c)). During phase ϕ_2, the capacitors denoted by $2C_D$ are discharged. In ϕ_1, the switches at the sources of either M_2 or M_6 are turned on, depending on the quantizer decision, D. The drain current of M_2 or M_6 (which should not go into the triode region during modulator operation) starts out at $V_{ref}/(R_D + 2R_s)$ and decays exponentially towards zero, as the capacitor $2C_D$ keeps charging. The decaying time constant of this current can be seen to be $(R_D+2R_s)C_D$. Replica switches are used in series with the sources of M_1 and M_3 to match the peak current/time constant of the assistant and feedback DACs. Capacitors C_b (approximately 5 pF) are required for high frequency bypass and are realized using MOS gates.

The reader might wonder why $2C_D$ and $0.5R_D$ are used to set the decay time constant in the assistant DAC, when the use of C_D and R_D could result in a smaller area. The reason for this choice is the following. Note that $\mathrm{M}_2/\mathrm{M}_6$ and $\mathrm{M}_4/\mathrm{M}_5$ should not

operate in the triode region. Using C_D and R_D would result in a higher(lower) quiescent potential at the gate of $\mathrm{M}_1(\mathrm{M}_3)$, thereby reducing the headroom available for $\mathrm{M}_2/\mathrm{M}_6$ and $\mathrm{M}_4/\mathrm{M}_5$.

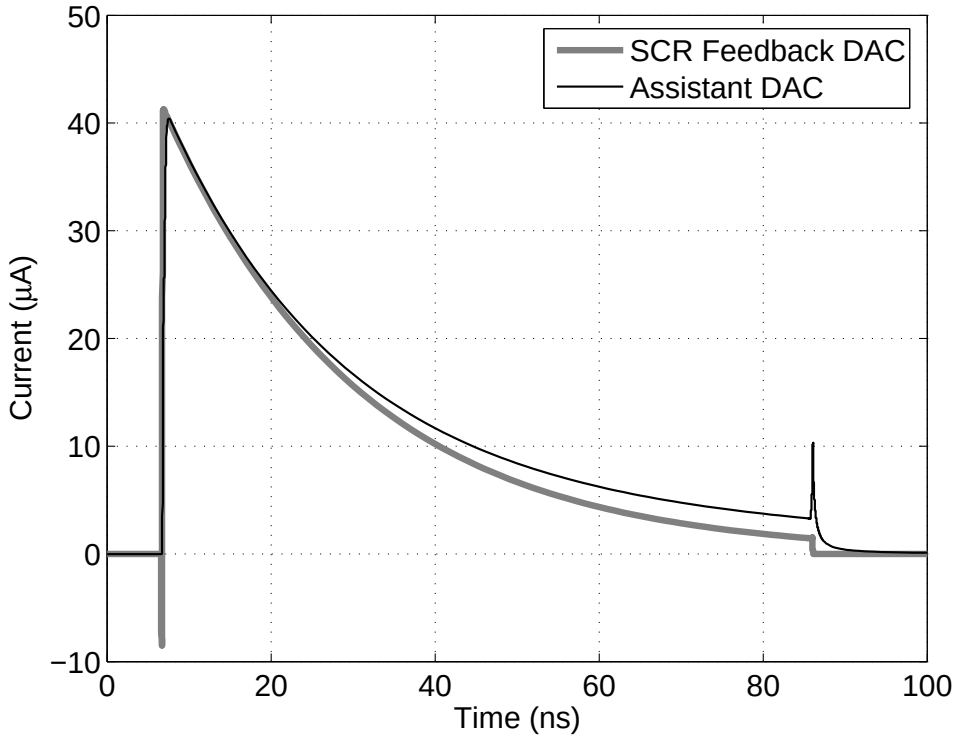


Figure 6.12: Pulse shapes of the SCR feedback and assistant DACs.

The differential component of the current pulses injected by the feedback and assistant DACs are shown in Fig. 6.12. Notice the good matching between the two pulses when the injected current is large. At the end of ϕ_2, the assistant pulse deviates from the feedback pulse by about 1.8 μA due to the decreasing g_m of $\mathrm{M}_2/\mathrm{M}_6$. This is not a problem as the difference is small. The current in the second stage of the opamp has been deliberately increased to 6 μA (from 4 μA in the NRZ case) per leg to be able to handle this mismatch.

The proposed architecture for the assistant DAC robustly addresses the issues that caused problems with the scheme of Fig. 6.10 in the following ways.

- *Initial current :* The peak current is desensitized with respect to error in the voltage at the gates of $\mathrm{M_1/M_3}$ due to resistive degeneration and the large $g_m R_D$ product in the initial portion of the DAC pulse. Charge sharing is less of an issue since the parasitic capacitors involved (source-bulk capacitance of the devices) are much smaller (compared to the gate-source capacitances involved in the generator of Fig. 6.10). The voltage drops across switches are no longer a problem as they are taken into account while setting the bias voltages at the gates of $\mathrm{M_1/M_3}$.
- *Time constant matching :* As shown in Fig. 6.12, the time constant of the assistant DAC pulse is matched well with that of the SCR feedback DAC waveform in the initial (high current) portion. As the current reduces, the time constant of the assistant pulse increases (due to the decreasing g_m of $\mathrm{M_2/M_6}$, $\mathrm{M_4/M_5}$).
- *Current in the off state :* When ϕ_1 is low, the sources of $\mathrm{M_2/M_6}$ and $\mathrm{M_4/M_5}$ are open, thus ensuring zero current.

6.4 Simulation results

Power reduction using the assisted opamp integrator : For estimating the benefit (power reduction) of using opamp assistance in the NRZ and SCR modulators, the modulators were simulated by replacing the assisted opamp integrator with a conventional active-RC integrator. The same two-stage feedforward topology (as in Fig. 6.4) was used for the opamp, keeping the first stage current fixed at 5 μA and scaling the second stage[4] to accommodate the current that the opamp needs to source/sink. The input to the modulator was chosen to be a -5 dBFS 6 kHz sinewave. In the NRZ case, the total current consumed by the assisted opamp integrator (opamp, assistant transconductor and DAC) is 27 μA and the simulated in-band SNDR is observed to be 107 dB[5]. When the assistant is removed and the quiescent current in the second stage of the opamp increased to 10 μA per leg (so that the current consumption is the same as the assisted integrator), the in-band SNDR is seen to be 89 dB. The PSD of both the modulators (with and without assistance, but same total power) is shown in Fig. 6.13. The inset in

[4] maintaining the same gate overdrives

[5] this is almost the simulated SNDR of a modulator with a linear loop filter.

the figure shows that the virtual ground node excursions are considerably smaller with opamp assistance.

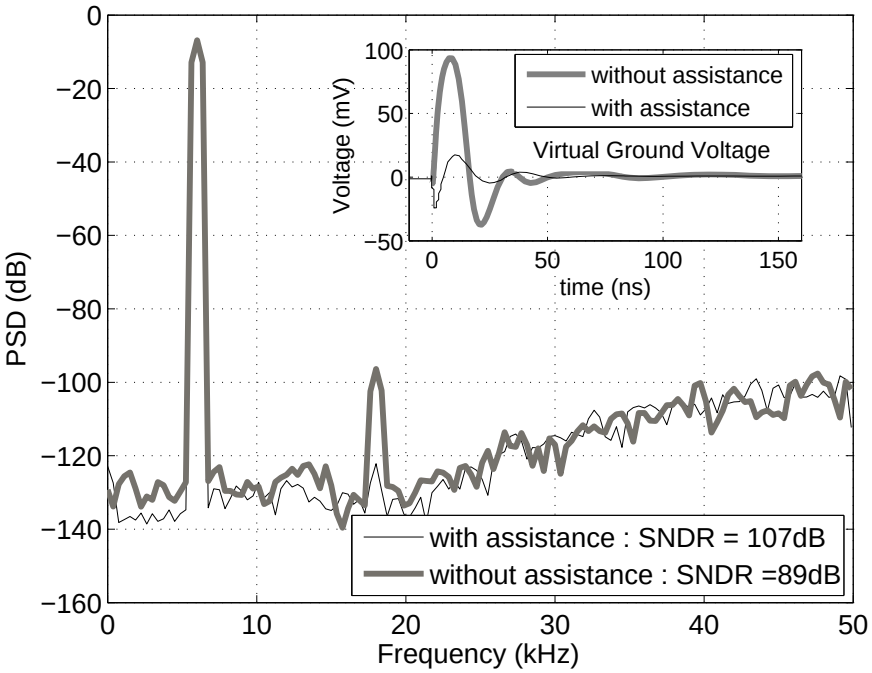


Figure 6.13: Low frequency PSD of NRZ modulators using conventional and assisted opamp integrators (with the same power consumption). The inset shows the voltages at the virtual ground node in both cases.

Further, to determine the improvement possible with the conventional integrator with enhanced power dissipation, the performance of the modulator was observed by progressively increasing the second stage current of the opamp. The observed results are as given in Tab. 6.1. The second row of the table corresponds to the usual practice of designing the opamp so that slewing is avoided (with some design margin). The maximum load current that the opamp needs to supply per leg is 13.5 μA. Using a design margin of 2.5 μA per leg in the second stage, a first stage current of 5 μA and CMFB/bias current of 2 μA (the same as in the assisted case), the resulting current consumption is 39 μA. Still, the in-band SNDR is not quite as high as in the assisted modulator, which clearly shows the power advantage obtained with opamp assistance. The

Total Current (μA)	Second Stage Current (μA)	In-band SNDR (dB)
27	20	89
39	32	102
47	40	103.5

Table 6.1: Performance of a CTDSM using a conventional integrator (NRZ DAC) as a function of current consumption.

same experiments were repeated with the SCR DAC based modulator as well. Fig. 6.14 compares the PSD of modulators with and without assistance, for the same current consumption in both cases (about 35 μA). It can be seen that using assistance improves the SNDR by about 23.5 dB. Table 6.2 gives the simulated performance of the SCR DAC based modulator (without assistance) as a function of the total bias current. The third row represents the situation where currents in the second stage of the opamp are chosen to avoid slewing.

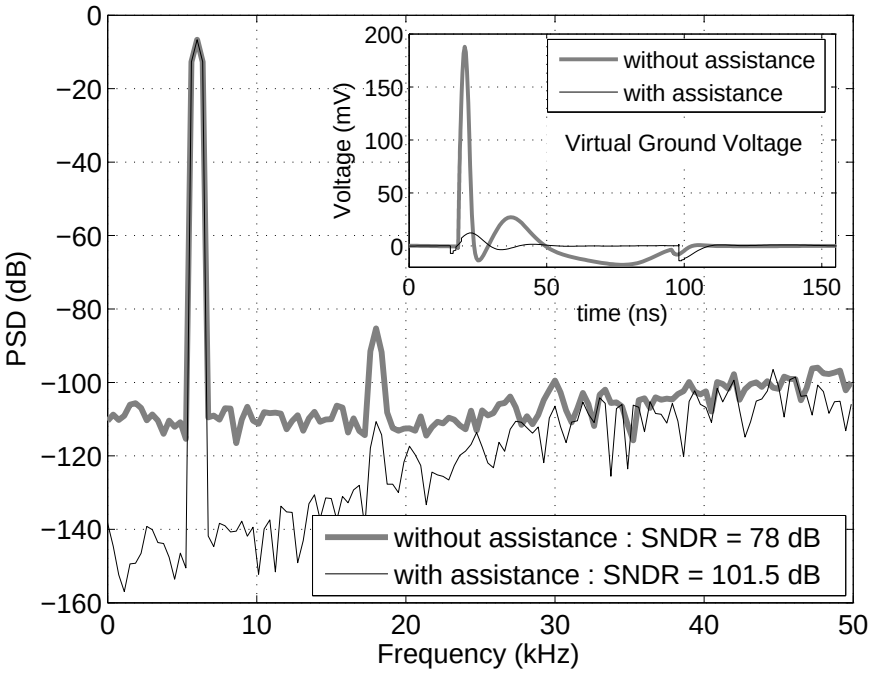


Figure 6.14: Low frequency PSD of SCR modulators using conventional and assisted opamp integrators (with the same power consumption). The inset shows the voltages at the virtual ground node in both cases.

In short, the simulation results shown above prove that the assisted opamp integrator

Total Current (μA)	Second Stage Current (μA)	In-band SNDR (dB)
35	28	78
67	60	98
113	106	102

Table 6.2: Performance of a CTDSM using a conventional integrator (SCR DAC) as a function of current consumption.

technique can reduce the power consumed in the first integrator. An added advantage of using an assisted opamp integrator is that, one can bypass the "trial and error" approach usually employed during the design of the first opamp in a CTDSM. The assisted opamp technique provides a simple, straight-forward and power efficient design methodology.

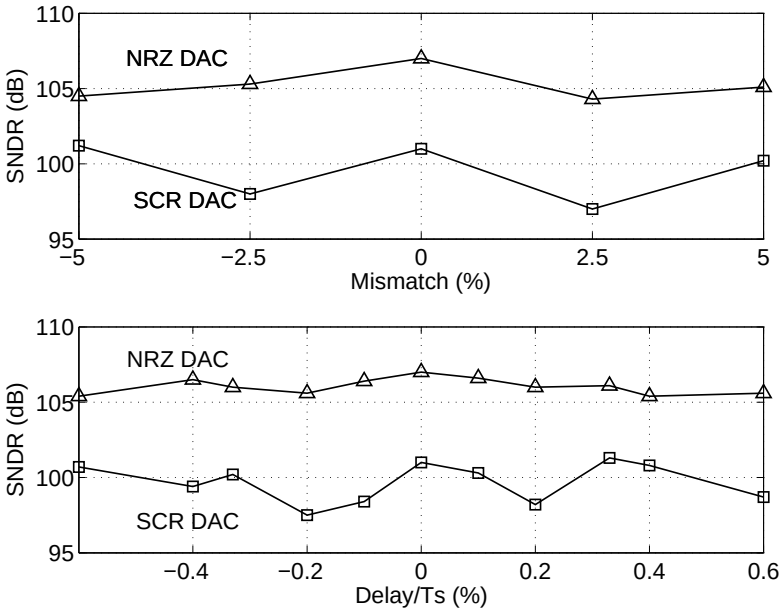


Figure 6.15: Simulated performance of modulators employing assisted integrators with NRZ and SCR DACs, with $\pm 5\%$ mismatch and ± 1 ns (0.6 %Ts) timing skew between feedback and assistant DACs.

Mismatch and timing errors : Since the technique of opamp assistance is based on cancellation of the swing at the virtual ground of the opamp, a potential problem is the sensitivity to mismatch and timing skew errors between the currents injected by the feedback and assistant DACs. To observe such effects in this design, simulations incorporating mismatch ($\pm 5\%$) and timing skew ($\pm 0.6\% T_s = 1\,\mathrm{ns}$) between the feedback

and feedforward DACs were run. Fig. 6.15 shows the results, indicating that the assisted opamp technique is robust in the face of such practical nonidealities.

6.5 Measurement results

The third order continuous-time delta-sigma modulators were fabricated in a 0.18 μm CMOS process through Europractice. The test setup used to measure the performance of the modulator is as shown in Fig. 6.16. Five hundred thousand samples from the modulator output, captured using an Logic analyzer were processed offline on a PC. Since the signal generator and the clock generator were not synchronized, a 32K Blackman-Harris window (rather than the Hann window used during simulations) was used to minimize the leakage effects during power spectral density estimation.

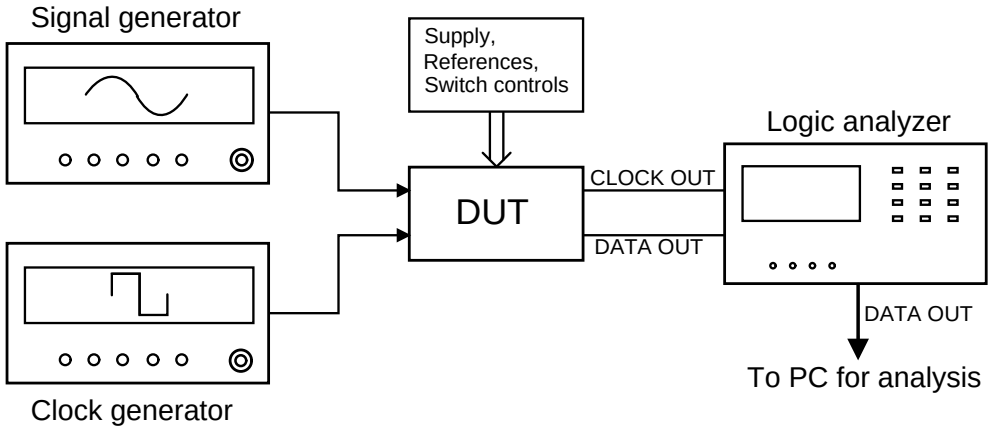


Figure 6.16: Block diagram representation of the setup used for measuring the performance of the modulators.

Fig. 6.17 shows the test board and layout screen shots of the active areas of the CTDSMs[6]. The NRZ and SCR designs occupy about 0.24 mm^2 and 0.26 mm^2 respectively. This is a third of the area occupied by the multi-bit modulator reported by Pavan *et al.* (2008), in the same process.

[6]The reader can refer to Appendix I, where the pin details of the modulator chip are provided.

Figure 6.17: Test board and chip layouts of the NRZ and SCR DAC based modulators.

Fig. 6.18 shows the measured SNR/SNDR of the modulators. The SNR is determined using a 15 kHz sinewave input, while a 6 kHz tone is used for SNDR measurement. The peak SNR/SNDR are 91 dB/88 dB and 90.3 dB/89.1 dB for the NRZ/SCR modulators. The measured dynamic ranges are 92.5 dB and 91.5 dB respectively.

The PSD plots of the modulator outputs are shown in Fig. 6.19. Part (a) of the figure shows the spectrum produced by the CTDSM with the NRZ DAC for a -2.2 dBFS 6 kHz sinewave input[7]. It can seen that the harmonics are about 95 dB below the fundamental and no non-harmonic tones are observed above the noise floor. Fig. 6.19(b) shows the output spectrum of the modulator with the SCR DAC for a -2.1 dBFS 6 kHz input. The third harmonic is seen to be about 97 dB below the fundamental.

Fig. 6.20 shows the measured rejection in the alias-band, (f_s - 24 kHz) to (f_s + 24 kHz). The simulated rejection (as computed from an AC response) is also shown for comparison and good agreement is seen. Around f_s, the measured rejection is around 90 dB, falling to 72 dB at the edge of the alias band. Fig. 6.21 shows the measured (and simulated) STF at out-of-band frequencies. Due to the nonlinear nature of the single-bit quantizer, the STF is a strong function of the amplitude of the sinusoidal input tone. The STF can be seen to peak significantly (15 dB) when a 50 mV input is used, but reduces to 5 dB when the amplitude is increased to 500 mV.

[7]This is the amplitude that results in the best SNDR.

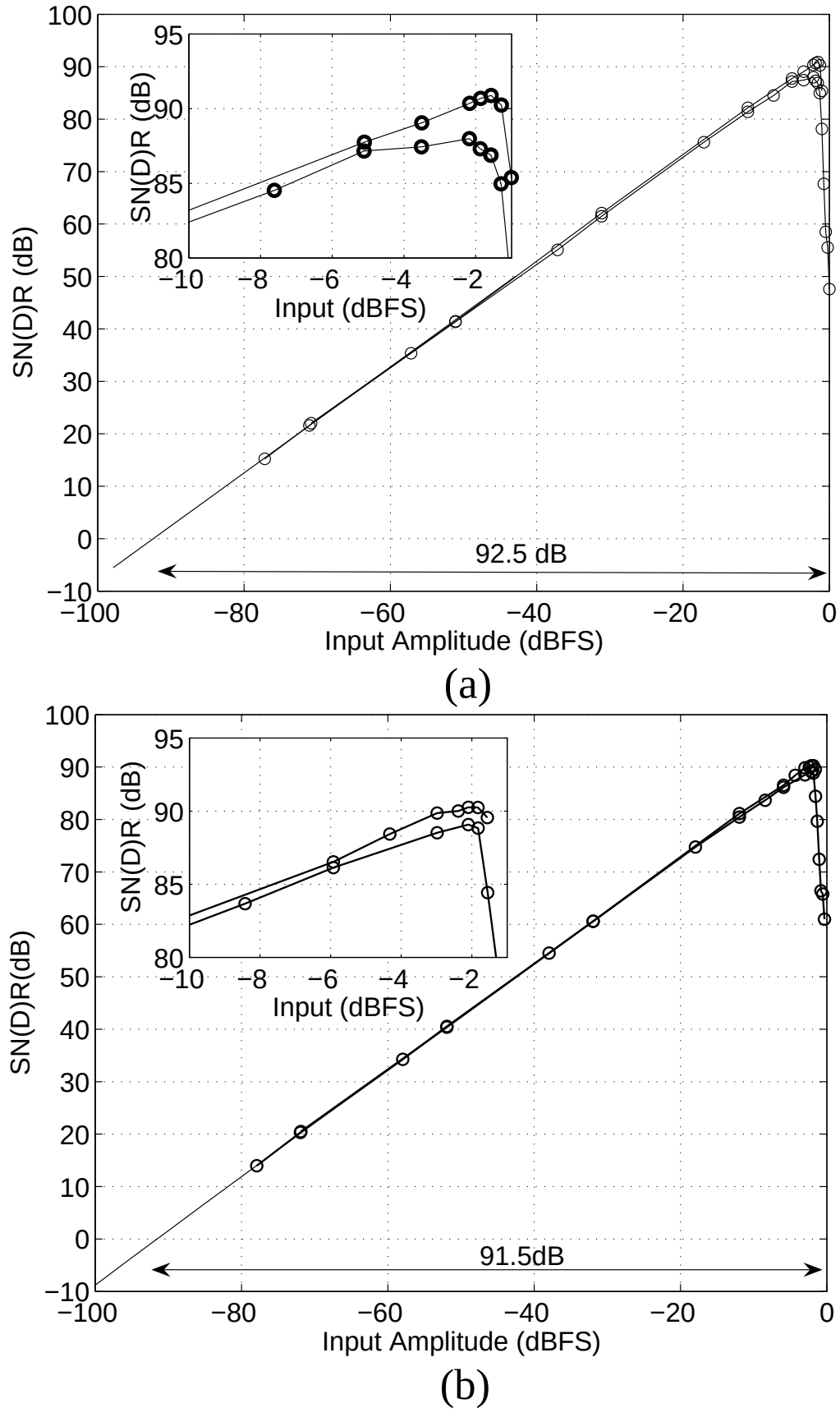


Figure 6.18: Measured SNR and SNDR for the modulator with (a) an NRZ DAC and (b) an SCR DAC. The dynamic ranges are 92.5 dB and 91.5 dB respectively.

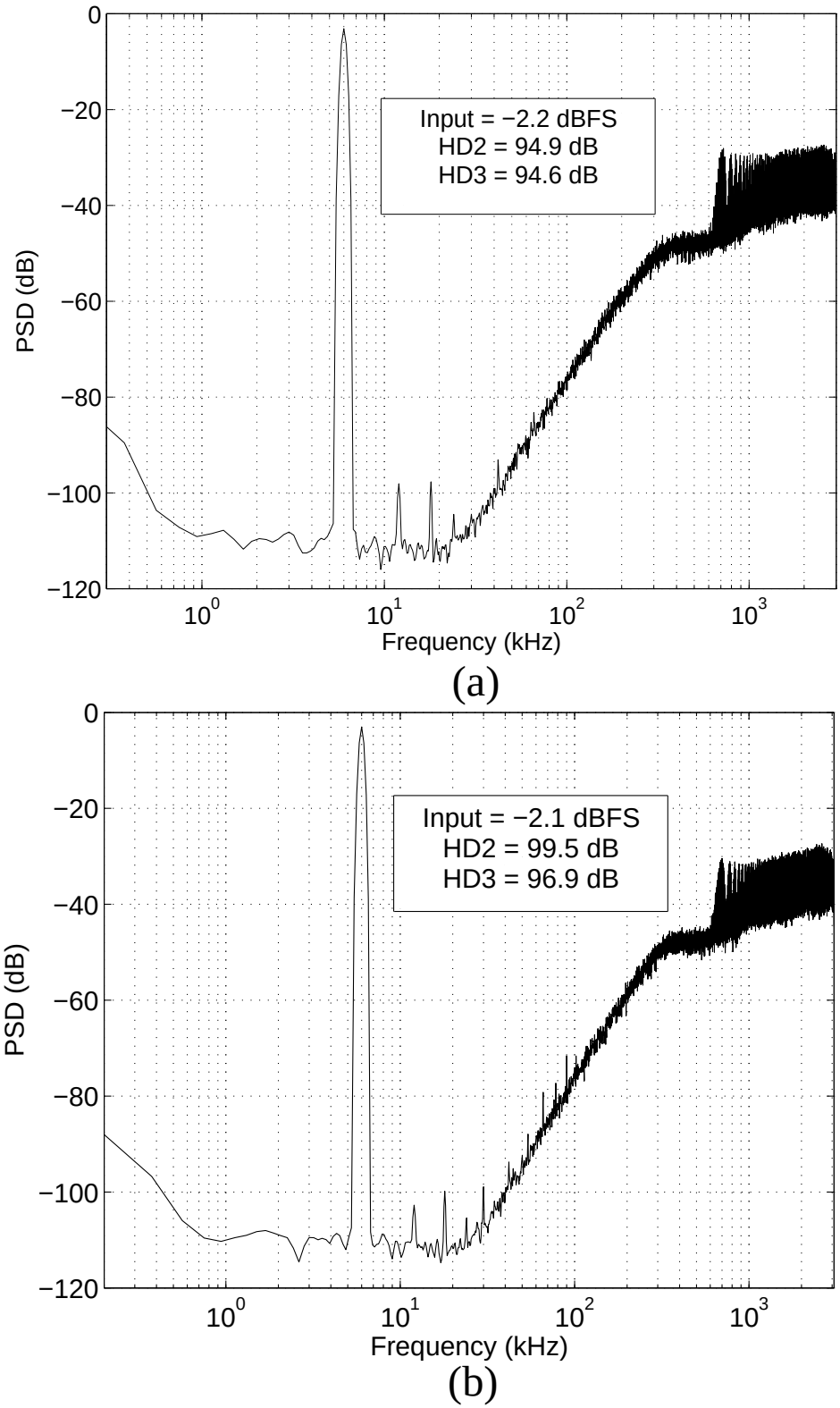


Figure 6.19: Modulator output spectrum for a 6 kHz tone (a) for an NRZ DAC based modulator with -2.2 dBFS input and (b) for an SCR DAC based modulator with -2.1 dBFS input.

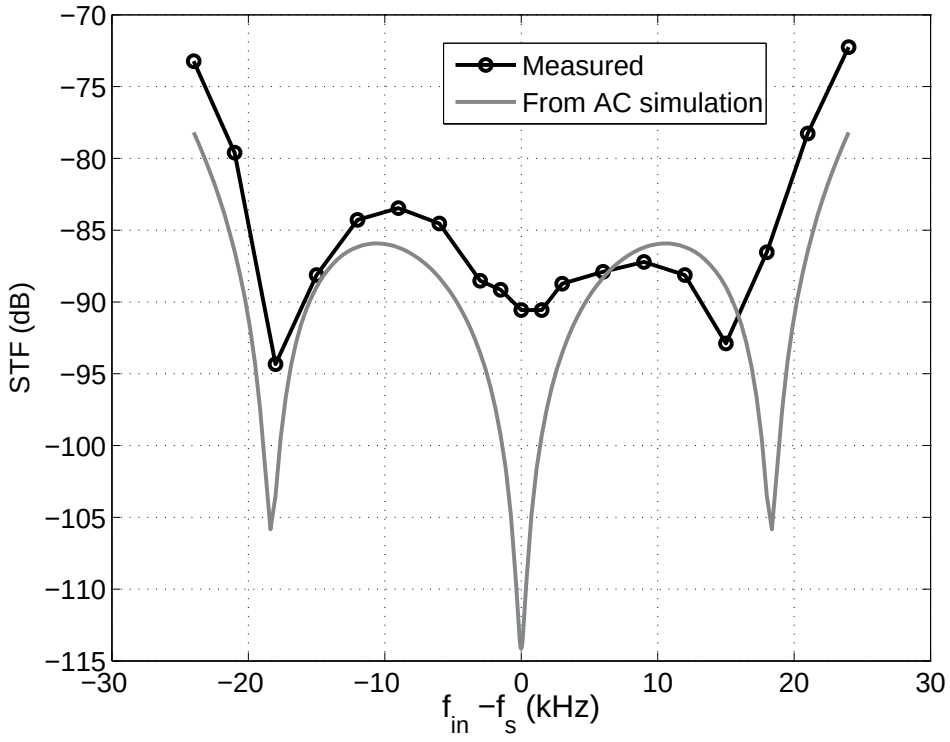


Figure 6.20: Measured and simulated rejection in the first alias band ((f_s - 24 kHz) to (f_s + 24 kHz))

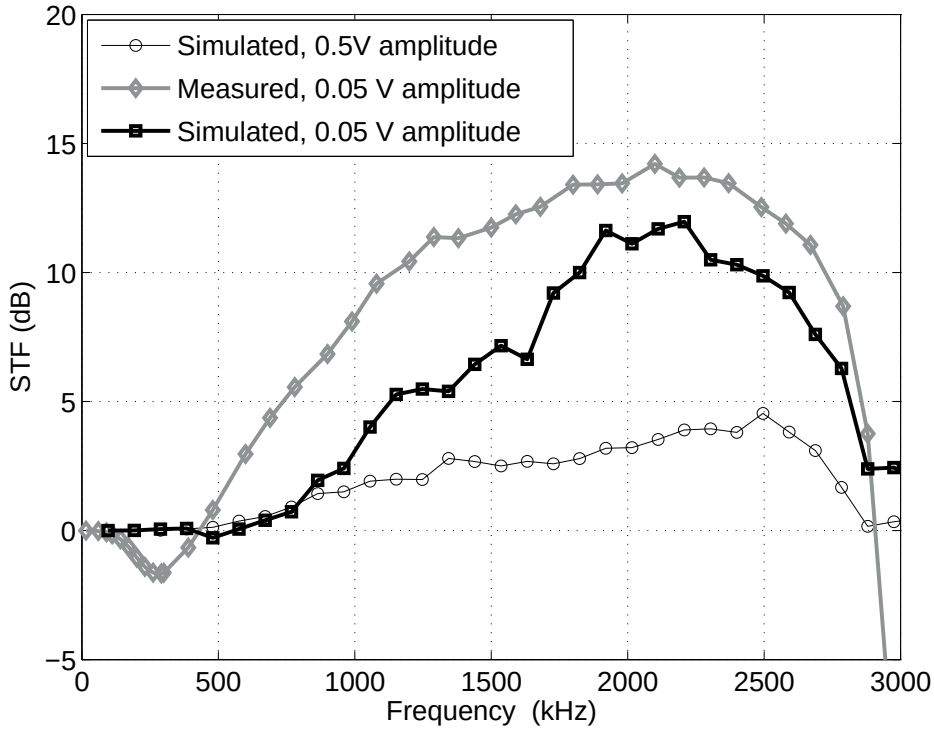


Figure 6.21: Measured and simulated STF for out-of-band frequencies. The magnitude of the STF depends significantly on the input amplitude.

6.5.1 Jitter sensitivity measurements

Since a controlled jitter injection setup was not available, the jitter performance of the two modulators was compared using an FM modulated sinusoidal clock source which can be expressed as

$$V_{clk} = A_{clk}\sin\left(2\pi f_s t + \frac{\Delta f}{f_m}\sin(2\pi f_m t)\right) \tag{6.1}$$

where f_s and f_m refer to the sampling and modulating frequencies and Δf denotes the peak frequency deviation. With such a modulated clock, the timing errors in the clock edges can be shown to be

$$\Delta T_{s1}[n] = \frac{\Delta f}{2\pi f_s f_m}\sin\big(2\pi f_m n T_s\big) \tag{6.2}$$

$$\Delta T_{s2}[n] = \frac{\Delta f}{2\pi f_s f_m}\sin\big(2\pi f_m (n+0.5) T_s\big) \tag{6.3}$$

For the modulator with NRZ DAC, the equivalent random error input sequence introduced as a result of the clock jitter is given as,

$$e_{j,NRZ}[n] = \frac{\Delta T_{s1}[n]}{T_s}\big(v[n] - v[n-1]\big) \tag{6.4}$$

$$= \frac{\Delta f}{2\pi f_m}\sin\big(2\pi f_m n T_s\big)\big(v[n] - v[n-1]\big) \tag{6.5}$$

If $f_m \gg f_s/OSR$, the timing uncertainty concentrated at a frequency f_m can fold back the out of band noise of $v[n] - v[n-1]$ into the in-band region. This can be shown to result in a noise, which is white in nature (Tao *et al.* (1999)). From the above equation, the timing uncertainty's tonal amplitude can be seen to be $\Delta f/(2\pi f_m)$. With a DAC reference of V_{ref}, the probability of the quantizer making a transition ($\pm 2V_{ref}$) can be

denoted by an activity factor, Λ. Thus, the in-band power of the (white) jitter induced noise is given by

$$\begin{aligned} J^2_{rms,NRZ} &= \frac{1}{2}\left(\frac{\Delta f}{2\pi f_m}\right)^2 \Lambda(2V_{ref})^2 \frac{1}{OSR} \quad \text{or,} \\ J_{rms,NRZ} &= 2V_{ref}\frac{\Delta f}{\pi f_m}\sqrt{\frac{\Lambda}{8\mathrm{OSR}}} \end{aligned} \tag{6.6}$$

With the NRZ design, measurements show that $\Lambda \approx 0.75$. Choosing f_m to be 2.5 MHz, the in-band noise of the modulator was measured as a function of Δf (with their inputs set to zero). Δf was varied upto 25 kHz, which corresponds to a peak-to-peak jitter of about 0.52 ns. Fig. 6.22 shows the measured in-band noise (denoted by the circles) along with the expected values as computed from eq. 6.6.

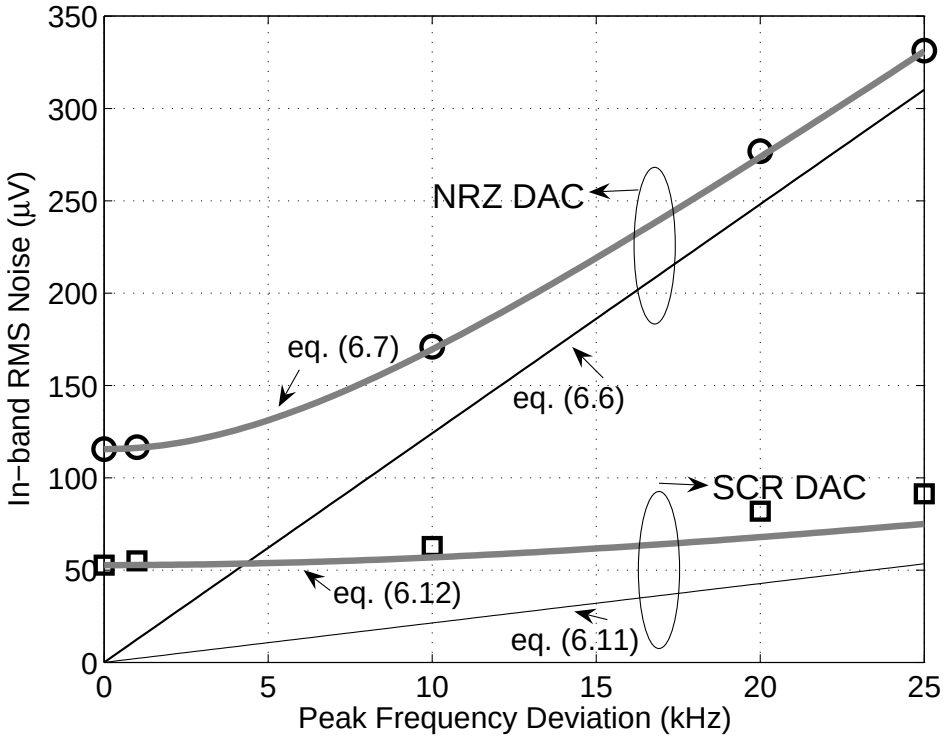


Figure 6.22: Measured jitter performance of the NRZ and SCR modulators, with an FM modulated sinewave clock source. The upper and lower curves in gray are calculations using eq. 6.7 and 6.12 respectively.

The reason for deviation of the measured noise from the expected jitter noise can be pointed out to the jitter of the clock source used for this experiment. Along with the thermal noise of the modulator, this explains the measured 116 μV RMS noise even when $\Delta f = 0$. This is several times higher than the noise seen with the (clean) square wave clock source, used for all the measurements reported previously.

However, when $\widehat{J}_{rms,NRZ}$ defined by

$$\widehat{J}_{rms,NRZ} = \sqrt{J_{rms,NRZ}^2 + (116\,\mu\mathrm{V})^2} \tag{6.7}$$

is plotted as a function of Δf (upper gray curve in Fig. 6.22), we can see that there is excellent matching between prediction and measurement. This makes sense because the thermal noise of the modulator, noise due to the intrinsic jitter of the clock source and that due to the FM tone are uncorrelated.

The jitter performance of an SCR modulator is expected to be better, when compared to an NRZ modulator. This is due to the fact that the current at the end of the clock phase ϕ_1 (denoted as I_1) in an SCR DAC will be smaller than the current of an NRZ DAC (I_{NRZ}). In the presence of an FM modulated clock (with timing errors in the clock edges given by eq. 6.2 and 6.3), the input-referred jitter noise sequence with an SCR DAC can be expressed as

$$\begin{aligned}
e_{j,SCR}[n] &= \left(\frac{\Delta T_{s1}[n]}{T_s} - \frac{\Delta T_{s2}[n]}{T_s}\right) v[n]\frac{I_1}{I_{NRZ}} && (6.8)\\
&= \frac{\Delta f}{2\pi f_m}\Big[\sin\big(2\pi f_m n T_s\big) - \sin\big(2\pi f_m (n+0.5) T_s\big)\Big] v[n]\frac{I_1}{I_{NRZ}} && (6.9)\\
&= \frac{\Delta f}{2\pi f_m}\Big[\big(1-\cos(\pi f_m T_s)\big)\sin\big(2\pi f_m n T_s\big) \\
&\qquad - \sin(\pi f_m T_s)\cos\big(2\pi f_m n T_s\big)\Big] v[n]\frac{I_1}{I_{NRZ}} && (6.10)
\end{aligned}$$

Similar to the case with the NRZ DAC, this results in a white noise if $f_m \gg f_s/OSR$. With the tonal amplitude of the timing uncertainty being given by $|1-\cos(\pi f_m T_s) - j\sin(\pi f_m T_s)|$, the RMS in-band noise with an SCR DAC can be approximately shown to be

$$J_{rms,SCR} \approx \beta V_{ref} \frac{\Delta f}{\pi f_m} \frac{I_1}{I_{NRZ}} \sqrt{\frac{1}{8\mathrm{OSR}}}, \qquad (f_m \gg f_s/OSR) \tag{6.11}$$

where $\beta = |1-\cos(\pi f_m T_s) - j\sin(\pi f_m T_s)|$. The measured in-band noise with jitter for the case of the SCR modulator is plotted by squares in Fig. 6.22. It can be seen that the noise is significantly smaller than that in the NRZ case, indicating superior performance of the SCR modulator when there is clock jitter. When

$$\widehat{J}_{rms,SCR} = \sqrt{J_{rms,SCR}^2 + (53\,\mu\mathrm{V})^2} \tag{6.12}$$

is plotted as a function of Δf, there is a disagreement between predictions and measurements for values around $\Delta f = 20\,\mathrm{kHz}$. A possible reason for this is the deviation of I_1 from the value predicted in simulation (1.5 μA), due to RC time constant variations.

Table 6.3: Summary of measured performance of the NRZ and SCR modulators.

	NRZ Modulator	**SCR Modulator**
Signal Bandwidth/Clock Rate	24 kHz / 6.144 MHz	24 kHz / 6.144 MHz
Quantizer Range	3.6 $\mathrm{V_{pp,diff}}$	3.6 $\mathrm{V_{pp,diff}}$
Input Swing for peak SNR	-1.6 dBFS	-2 dBFS
Dynamic Range/SNR/SNDR	92.5 dB/91 dB/88 dB	91.5 dB/90.3 dB/89.1 dB
Active Area	0.24 $\mathrm{mm^2}$	0.26 $\mathrm{mm^2}$
Process/Supply Voltage	0.18 μm CMOS/1.8 V	0.18 μm CMOS/1.8 V
Power Dissipation (Modulator + References)	110 μW	122 μW
Figure of Merit (SNDR)	112 fJ/level	109 fJ/level
Figure of Merit (Schreier)	175.9	174.4

A summary of the measured performance of the modulators is given in Table 6.3. The Figure of Merit (FOM) of the modulators is determined as (Gerfers *et al.* (2003))

$$FOM_{SNDR} = \frac{P}{2 \times f_b \times 2^{(SNDR-1.76)/6.02}} \tag{6.13}$$

where P and f_b denote the power dissipation and signal bandwidth respectively. Since the performance is dominated by thermal noise, it is also relevant to compute the FOM proposed by Schreier (Schreier and Temes (2005*b*)), which is given by

$$FOM_{DR} = \text{Dynamic Range}_{\text{dB}} + 10\log_{10}\left(\frac{f_b}{P}\right)$$

Note that FOM_{DR} is a logarithmic measure. The NRZ and SCR based modulators achieve FOM_{SNDR} of 112 fJ/level and 109 fJ/level and FOM_{DR} of 175.9 and 174.4 respectively. The FOMs are close to that achieved in the multi-bit design reported by Pavan *et al.* (2008), but with one-third the active area.

6.6 Comparison with other work

Table 6.4 compares the performance of several delta-sigma modulators presented in the literature with the designs implemented in this research. Power efficiency of the designs can be compared using the FOM provided. The modulators reported by Kim *et al.* (2008), Roh *et al.* (2008), Chae and Han (2009), Park *et al.* (2009) and Yao *et al.* (2004) are discrete-time designs. The designs of Gerfers *et al.* (2003) and Ortmanns *et al.* (2005) use single-bit multiple feedback continuous-time architecture and achieve SNDRs of 70 and 65 dB respectively. The latter uses an SCR feedback DAC. It appears that the linearity of these modulators is limited by the first integrator.

The ADC presented by Nguyen *et al.* (2005) employs a four-bit quantizer in the loop and has a hybrid continuous-time and switched-capacitor loop filter. Apart from the modulator, the chip also includes RC time constant tuning loops and a "clock clean-up" loop that makes the performance insensitive to clock jitter. Dorrer *et al.* (2006) presents a design in a 65 nm CMOS process. It employs a "tracking ADC" to reduce the number of comparators in the 4-bit quantizer. Due to the high OSR, the power efficiency is low.

The modulator reported by van der Zwan and Dijkmans (1996) employs a single-bit modulator and is intended for voice applications. Pun *et al.* (2007) have shown a single-bit design in a 0.5 V supply, yielding 74 dB SNDR. The performance of the modulator is limited due to single-bit operation and the linearity of the first integrator. The multi-bit continuous-time design reported by Pavan *et al.* (2008) consumes a power of 121 μW (including reference buffers) and achieves a FOM of 65 fJ/level.

It can be seen that the NRZ modulator described in this chapter achieves a FOM which is better than that of the single-bit CTDSMs reported earlier. Further, it achieves an efficiency that approaches that of the modulator of Pavan *et al.* (2008)[8]. The SCR DAC based modulator is also power efficient and achieves very low distortion, because of the use of the assisted opamp technique.

[8] which uses a 4-bit quantizer and occupies thrice the area

Table 6.4: Comparison with other Delta-Sigma modulators

Reference	**B/W (kHz)/ OSR**	**Type (DT/ CT)**	**DR/SNR/ SNDR (dB)**	**Power (μW)**	**Tech. (μm)**	**Supply (V)**	**FOM$_{SNDR}$ (pJ/level)**	**FOM$_{DR}$ (dB)**
Kim *et al.* (2008)	24/128	DT	92/91/89	1500	0.13	0.9	1.36	164
Yao *et al.* (2004)	20/100	DT	88/85/81	140	0.09	1	0.38	169.5
Roh *et al.* (2008)	20/50	DT	83/82.2/73.1	60	0.13	0.9	0.406	168.2
Chae and Han (2009)	20/100	DT	85/84/81	36	0.18	0.7	0.098	172.4
Park *et al.* (2009)	25/100	DT	100/100/95	870	0.18	0.7	0.38	174.6
Gerfers *et al.* (2003)	25/48	CT	80/73/70	135	0.5	1.5	1.04	162.7
Ortmanns *et al.* (2005)	25/48	CT	81/66/66	250	0.5	1.5	3.07	161
Nguyen *et al.* (2005)	20/128	CT	106/106/98	18000	0.35	3.3	6.93	166.5
Dorrer *et al.* (2006)	20/300	CT	95/77/ 74	2200	0.065	1.2	13.43	164.6
van der Zwan (1996)	3.4/64	CT	80/–/70	210	0.5	2.2	11.95	152.1
Pun *et al.* (2007)	25/64	CT	74/74/74	300	0.18	0.5	1.46	153.2
Pavan *et al.* (2008)	24/64	CT	93.5/92.5/90.8	121	0.18	1.8	0.089	176.5
This work (NRZ)	24/128	CT	92.5/91/88	110	0.18	1.8	0.112	175.9
This work (SCR)	24/128	CT	91.5/90.3/89.1	122	0.18	1.8	0.109	174.4

CHAPTER 7

CONCLUSION

In this research, the issue of integrator nonlinearity in continuous-time delta-sigma modulators has been investigated. Several areas which had not been attended to in the past, have been studied. The contributions and observations borne out of this research can be summarized as below.

7.1 Summary of the contributions and observations

The performance of multi-bit CTDSMs is observed to degrade due to integrator nonlinearity through an increase in the in-band noise floor. In order to analyze such performance degradations, suitable models were developed for nonlinear integrators that are employed as the first integrator of a CTDSM. Nonlinearity in the active-RC integrators (which are closed loop systems) was analyzed by using Bussgang's method of current injection. With the input to the integrator being a sampled and held version of the shaped quantization noise, two-stage active-RC integrators were analyzed using a simple, intuitive approach. From these models, it was inferred that active-RC integrators built with feedforward compensated opamps are the most linear among various integrator topologies.

Using these models in a multi-bit modulator with CIFF loop filter, the effect of integrator nonlinearity was perceived as introducing an additional noise sequence at the input of a linear modulator. This noise sequence was shown to be a nonlinear function

of the shaped quantization noise. With the assumption of Gaussian distribution for the shaped quantization noise, analytical expression were thus derived, quantifying the increase in the in-band noise of multi-bit modulators. These expressions clearly bring out the influence of various parameters like the NTF of the modulator (OBG), the number of quantizer levels and the nonlinearity of the integrator (β), on the performance of a multi-bit CTDSM.

When the analysis was extended to modulators employing RZ feedback DACs, several useful observations were obtained. With an RZ DAC, the in-band noise due to integrator nonlinearity is 12 dB higher than that observed with an NRZ DAC. This can be attributed to the fact that the height of the RZ pulse is twice that of an equivalent NRZ pulse. Further, with an RZ DAC, the in-band noise was found to increase as a function of the input signal amplitude, thus leading to even higher degradations in the modulator performance.

Results from macromodel simulations with various integrator topologies and different number of quantizer levels proved the accuracy of the derived models. However, with active-RC integrators having feedforward compensated opamps, the in-band noise was shown to be under-estimated using the analytical expressions. This was because the initial transients at the virtual ground node were neglected while deriving the nonlinear models for such integrators. This was duly addressed by modeling the initial transients with equivalent sampled and held components. This resulted in expressions for the in-band noise due to integrator nonlinearity, which can be quickly computed through numerical calculations using tools like MATLAB.

The analysis of integrator nonlinearity was also extended to study the harmonic distortion observed in single-bit modulators. Similar to the multi-bit modulator, the initial transients were shown to increase the harmonic distortion in the modulator output. This was attributed mainly to the interference between consecutive output samples of the modulator, with an NRZ feedback DAC. Since such interferences do not occur for modulators employing RZ DAC, the harmonic distortion was shown to be independent of the initial transients.

Inspite of the better performance with an RZ DAC, it was realized that NRZ architectures are preferred for their relaxed slew-rate and jitter constraints. The conventional way of addressing the issue of integrator nonlinearity in NRZ modulators was seen to increase the power consumed in the first integrator (opamp) in trying to improve the bandwidth and linearity of the opamp. The motivation to find an alternate power efficient solution directed us to the 'assisted opamp' technique. This was based on the intuition that if the opamp is assisted in sourcing/sinking the steps of DAC current, the swings at all its transconductor inputs can be minimized. By this way, the effect of nonlinearity of the transconductors can be reduced, thus resulting in a low distortion modulator performance. More importantly, it was shown that this technique can improve the linearity of the integrator with reduced power consumption and also without affecting the noise performance.

Finally, the design of audio frequency third order CTDSMs (with NRZ and SCR feedback DACs) employing the 'assisted opamp' technique was discussed. Simulation and experimental results clearly proved the efficacy of this technique in reducing the power consumption of the first opamp of a single-bit CTDSM. With the help

of opamp assistance, the NRZ/SCR modulators have achieved a dynamic range of 92.5 dB/91.5 dB while dissipating 110 μW/122 μW from a 1.8 V supply. The power efficiency achieved by the modulators was shown to be comparable to the best multi-bit modulators reported with similar specifications.

In short, an attempt has been made in this research to answer several unanswered questions related to the effects of integrator nonlinearity in CTDSMs. The proposed analytical models quantifying these effects can aid the design of low power, high resolution CTDSMs. Further, the proposed assisted opamp technique can help in reducing the power consumption of the loop filter. However, it should be mentioned that there is still a lot of scope for investigation in this interesting and important topic of integrator nonlinearity. The need for investigation can first be pointed out to the following limitations of the analysis carried out in this research.

- In case of multi-bit modulators, numerical computations are required with the proposed model to accurately quantify the effect of initial transients. Though the numerical method can quickly estimate the modulator performance, they provide comparatively lesser intuition than analytical expressions. While some useful insights were obtained with an approximate analytical expression in this work, further insights on the effect of initial transients can be achieved, if one can arrive at an accurate analytical expression.
- Similarly, in case of single-bit modulators with NRZ feedback DAC, the distortion performance is observed to be dependent on the input signal amplitude (through the factor γ_{in}) due to the presence of initial transients. In essence, mixing of the modulator output samples with its delayed version was shown as the cause for such a dependency. With little intuition on the magnitude of the second harmonic generated through the mixing, ideal modulator simulations were used in this research to compute γ_{in}. It would be useful and interesting to determine the influence of the input signal amplitude and/or the NTF of the modulator on this factor.

In addition, the analysis carried out in this research can be extended to several other architectures as explained below.

7.2 Suggestions for future work

A.) Effect of nonlinearity in subsequent integrators of a CIFB loop filter : The effect of first integrator's nonlinearity in a CIFB loop filter can be analyzed in the same manner as in a CIFF loop filter. While the nonlinearity of the subsequent integrators are not critical in a CIFF loop filter, it may not be the case in a CIFB architecture (Leuciuc (2001)). Therefore, there arises a need to analyze the effect of nonlinearity in these integrators in a CIFB loop filter. The same approach as discussed for the first integrator can be adopted for the analysis. However, a notable difference in these cases lies in the nature of the input to the integrators. Unlike the first integrator, the input to the other integrators is not only a sampled and held version of the shaped quantization noise. The input also includes filtered versions of this noise. This complicates the computations involved and hence, is a challenging task to quantify the effect of nonlinearity in subsequent integrators.

B.) Effect of nonlinearity in modulators with SCR feedback DAC : An even more challenging task is to analyze the effect of integrator nonlinearity in modulators employing SCR feedback DAC. This is because the exponential decaying current from the SCR DAC complicates the analysis unlike the NRZ/RZ case, where the output of the DAC is constant in a given clock period.

C.) Effect of timing skews when opamp assistance is employed in high-speed modulators: The technique of opamp assistance can be suitably applied for higher modulator speeds as well. However, the timing skew between the assistant and the feedback DAC current becomes increasingly crucial as the modulator speed is increased. This can be modeled as an error current injected at the opamp output, containing mostly of high fre-

quency components. This can lead to excursions at the virtual ground node and thereby distortion due to the opamp nonlinearity. Quantitative estimation of the modulator performance in such cases is highly dependent on the specifics of the opamp, in particular on its dynamics. It is an interesting area to study and to characterize the performance degradations observed with various opamp topologies.

In summary, several avenues are open to extend the approaches and results developed in this research and to quantify the effect of integrator nonlinearity in varied architectures of continuous-time delta-sigma modulators.

APPENDIX A

Analysis of nonlinear systems using Bussgang's method of current injection

Nonlinear differential equations are often encountered during circuit analysis due to the presence of nonlinear elements in the network. The Volterra series approach can be employed to obtain the circuit solutions in such cases. Using this approach, Bussgang *et al.* (1974) has provided a technique for solving circuits with weak nonlinearity. The technique recursively solves the linear equations of the network to obtain an approximate solution.

The steps involved in this method can be described with a help of a weakly nonlinear single-stage active-RC integrator as shown in Fig. A.1.

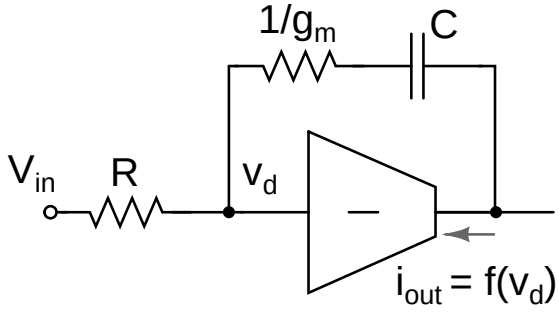


Figure A.1: A nonlinear active-RC integrator.

The transconductor is assumed to be weakly nonlinear[1] with

$$i_{out} = f(v_d) \tag{A.1}$$

$$= g_1 v_d + g_2 v_d^2 + g_3 v_d^3 + \cdots \tag{A.2}$$

[1]With weak nonlinearity, it can be assumed that $g_1 v_d \gg g_2 v_d^2 + g_3 v_d^3 + \cdots$.

Writing the nodal equation at the input of the transconductor,

$$\frac{V_{in} - v_d}{R} = i_{out} = f(v_d) \quad \text{or,} \tag{A.3}$$

$$\frac{V_{in}}{R} = \frac{v_d}{R} + f(v_d) \tag{A.4}$$

With a linear integrator, the above equation would have reduced to

$$\frac{V_{in}}{R} = \frac{(1 + g_1 R)v_d}{R} \tag{A.5}$$

The nonlinear equation given in eq. A.4 can be solved using Bussgang's method of current injection, by iteratively solving the above linear equation in the following manner,

- First, the linear solution from the above equation is determined as $v_{d,1}$.

$$v_{d,1} = \frac{V_{in}}{1 + g_1 R} \tag{A.6}$$

- Next, considering the second order nonlinearity, a nonlinear current, $i_{inj2} = -g_2 v_{d,1}^2$ is injected at the output of the transconductor as shown in Fig. A.2, with zero input ($V_{in} = 0$).

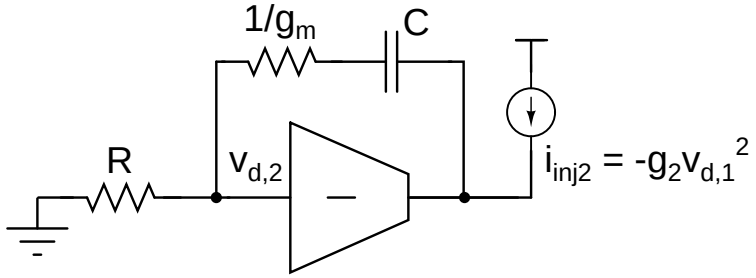


Figure A.2: Injection of the nonlinear currenti_{inj2} to a linear active-RC integrator.

Denoting $v_{d,2}$ as the solution obtained with this circuit,

$$v_{d,2} = \frac{R}{1 + g_1 R} i_{inj2} = \frac{-g_2 v_{d,1}^2 R}{1 + g_1 R} \tag{A.7}$$

$$= \frac{-g_2 V_{in}^2 R}{(1 + g_1 R)^3} \tag{A.8}$$

If one assumes $g_i = 0$ (for $i \geq 3$), the solution for v_d can be approximated as the sum of the linear solution (eq.A.6) and the nonlinear solution as obtained above.

$$v_d \approx v_{d,1} + v_{d,2} \tag{A.9}$$

- With higher order nonlinearities, the above procedure of injecting a nonlinear current can be repeated to determine the nonlinear part of the solution at v_d. In each step, the injected current is found using the solution obtained for v_d in the previous step.

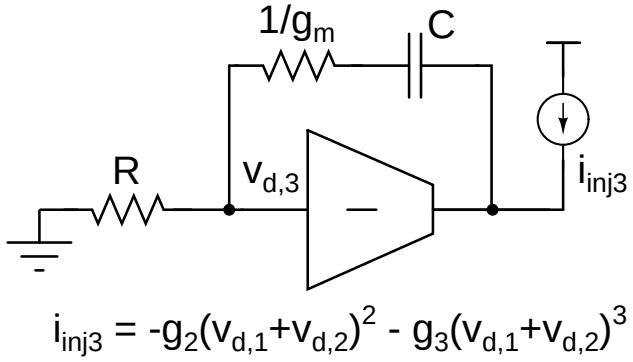


Figure A.3: Injection of the nonlinear currenti_{inj3} to a linear active-RC integrator.

For instance, with the third order nonlinearity, the nonlinear current $i_{inj3} = -g_2(v_{d,1} + v_{d,2})^2 - g_3(v_{d,1} + v_{d,2})^3$ is injected as shown in Fig. A.3. The solution thus obtained will be

$$\begin{aligned} v_{d,3} &= \frac{-g_2R(v_{d,1} + v_{d,2})^2}{1 + g_1R} - \frac{g_3R(v_{d,1} + v_{d,2})^3}{1 + g_1R} \\ &= \frac{-g_2R(v_{d,1}^2 + v_{d,2}^2 + 2v_{d,1}v_{d,2})}{1 + g_1R} - \frac{g_3R(v_{d,1}^3 + v_{d,2}^3 + 3v_{d,1}^2v_{d,2} + 3v_{d,2}^2v_{d,1})}{1 + g_1R} \end{aligned} \tag{A.10}$$

With weak nonlinearity and $g_1R \gg 1$, the solution can be approximated to

$$v_{d,3} \approx \frac{-g_2R(v_{d,1}^2 + 2v_{d,1}v_{d,2})}{1 + g_1R} - \frac{g_3Rv_{d,1}^3}{1 + g_1R} \tag{A.11}$$

Thus, the solution for v_d can now be obtained as

$$v_d \approx v_{d,1} + v_{d,3} \tag{A.12}$$

$$\approx v_{d,1} - \frac{g_2R(v_{d,1}^2 + 2v_{d,1}v_{d,2})}{1 + g_1R} - \frac{g_3Rv_{d,1}^3}{1 + g_1R} \tag{A.13}$$

The solution can so be easily extended to include higher order nonlinearities, by following the same procedure in an interative fashion.

It must be noted that with a differential implementation for the active-RC integrator,

even order nonlinearities do not exist. Therefore, the above procedure can be suitably simplified to include only the odd order nonlinearities. Though the procedure was discussed for a nonlinear single-stage active-RC integrator, it can be extended to any weakly nonlinear network. In essence, the key to solve a given nonlinear system is to recursively determine the solution of its linear equivalent, by injecting the necessary nonlinear currents at each iteration.

APPENDIX B

Comparison of Miller and feedforward compensated opamp based active-RC integrators

The choice of opamp architecture in an active-RC integrator can be made based on several criteria. It was shown in Chapter 2 that the smaller value of c_p in a feedforward opamp results in better linearity than a corresponding integrator built with Miller opamp. Though an useful observation from a design point of view, it is insufficient to comment on the power efficiencies of the two architectures. For a proper comparison of the integrators, an additional design parameter needs to be taken into account. In this discussion, we intend to compare the linearity (and the power efficiency) achieved with the two integrator architectures, when designed to have the same phase margin.

B.1 Active-RC integrator with a two-stage Miller compensated opamp

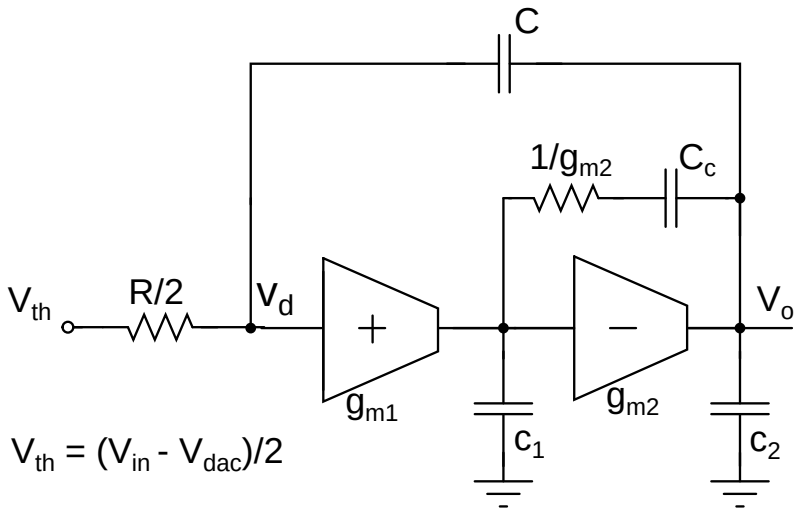


Figure B.1: Active-RC integrator with a two-stage Miller compensated opamp

Consider an active-RC integrator with a two-stage Miller compensated opamp forming the input stage of a CTDSM as shown in Fig. B.1. Stability analysis for the integra-

tor can be performed using the open loop configuration[1] as shown in Fig. B.2.

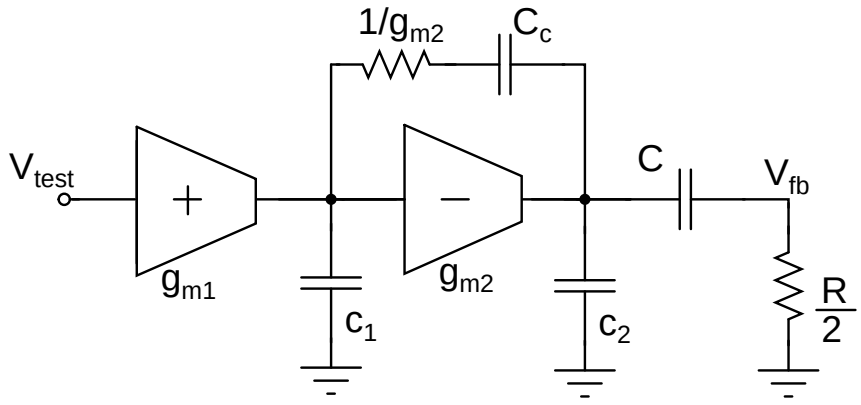


Figure B.2: Open loop configuration of the active-RC integrator with two-stage Miller compensated opamp

Note that a practical active-RC integrator is designed in a manner such that the parasitic pole (p_2) of the opamp lies much beyond the unity gain bandwidth of the opamp (ω_u). This in turn is designed to be much larger than the unity gain frequency of the integrator (ω_o). With reference to the circuit shown in Fig. B.1, we have

$$\frac{g_{m2}}{c_1 + c_2} \gg \frac{g_{m1}}{C_c} \gg \frac{2}{RC} \tag{B.1}$$

With the feedback network presenting a resistive load to the opamp for frequencies greater than $2/(RC)$, observe that the unity gain frequency of the (open) loop, $\omega_{u,loop}$ will be same as $\omega_u = g_{m1}/C_c$. Denoting the parasitic pole of the open loop system by $p_{2,loop}$, it is given by

$$p_{2,loop} = \frac{g_{m2} + 2/R}{c_1 + c_2} \tag{B.2}$$

Since the open loop network is predominantly a first order system, whose high frequency parasitic pole ($p_{2,loop}$) lies much beyond its unity gain frequency ($\omega_{u,loop}$), the

[1]The capacitance at the virtual ground node of the opamp has been neglected in this analysis. It is assumed that the virtual ground node capacitance is smaller when compared to the integrating capacitor.

phase margin of the active-RC integrator can be determined as

$$PM = 90^\circ - \tan^{-1}\left(\frac{\omega_{u,loop}}{p_{2,loop}}\right) \quad \text{(B.3)}$$

$$= 90^\circ - \tan^{-1}\left(\frac{g_{m1}(c_1+c_2)}{(g_{m2}+2/R)C_c}\right) \quad \text{(B.4)}$$

$$\approx 90^\circ - \tan^{-1}\left(\frac{g_{m1}(c_1+c_2)}{g_{m2}C_c}\right), \quad (\text{with } g_{m2} \gg 2/R\,) \quad \text{(B.5)}$$

B.2 Active-RC integrator with a feedforward compensated opamp

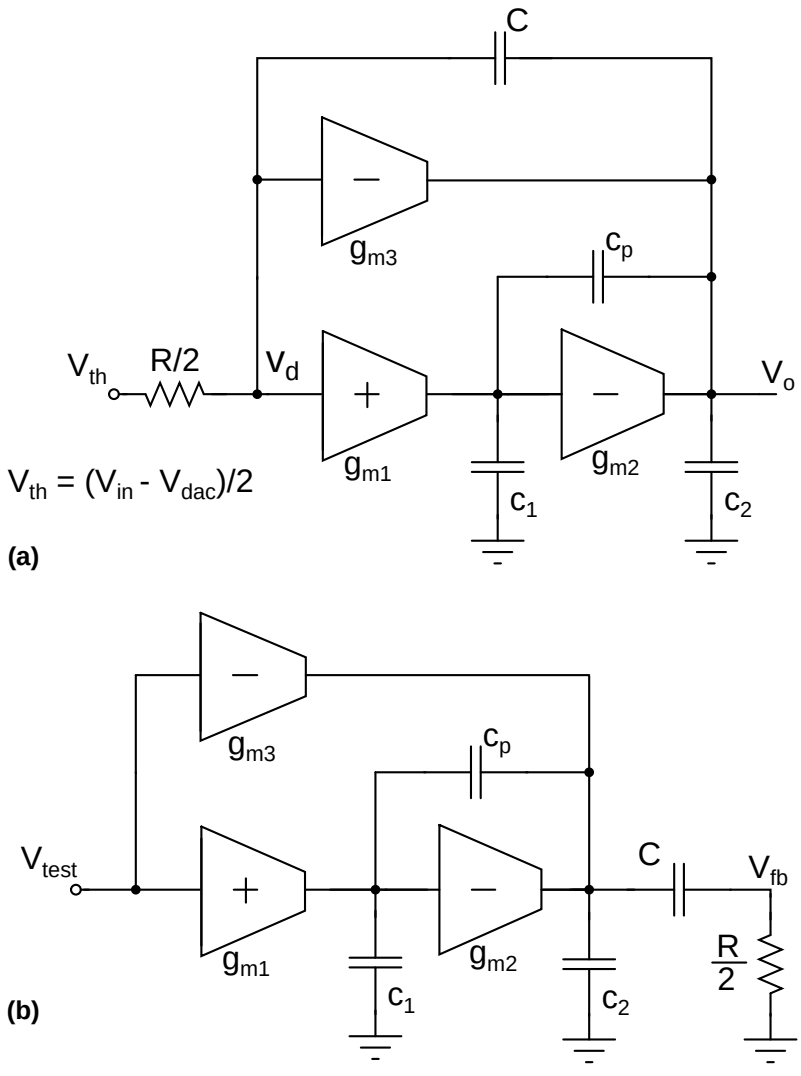


Figure B.3: (a) Active-RC integrator with a feedforward compensated opamp and its (b) open loop configuration used for stability analysis

Consider an active-RC integrator built with a feedforward compensated opamp as

shown in Fig. B.3(a). The open loop configuration used for stability analysis is shown in Fig. B.3(b). Similar to the case of Miller opamp, it can be assumed that the unity gain frequency of the integrator (ω_u) or that of the open loop ($\omega_{u,loop}$) is much higher than $\omega_0 = 2/(RC)$. Assuming the feedforward zero frequency ($\omega_{z,loop}$) to occur well before $\omega_{u,loop}$, the phase margin of the integrator can be expressed as

$$\begin{aligned} PM &= \tan^{-1}\left(\frac{\omega_{u,loop}}{\omega_{z,loop}}\right) && \text{(B.6)} \\ &\approx \tan^{-1}\left(\frac{g_{m3}/c_2}{g_{m1}g_{m2}/(g_{m3}c_1)}\right) && \text{(B.7)} \\ &= \tan^{-1}\left(\frac{g_{m3}^2 c_1}{g_{m1}g_{m2}c_2}\right) && \text{(B.8)} \end{aligned}$$

Equating eq. B.5 and B.8, the condition for both the integrator architectures to have the same phase margin is,

$$90^\circ - \tan^{-1}\left(\frac{g_{m1}(c_1+c_2)}{g_{m2}C_c}\right) = \tan^{-1}\left(\frac{g_{m3}^2 c_1}{g_{m1}g_{m2}c_2}\right) \quad \text{(B.9)}$$

Using the relation, $\tan^{-1}(\theta) + \tan^{-1}(1/\theta) = 90^\circ$, we have

$$g_{m3}^2 = g_{m2}^2 \frac{c_2 C_c}{c_1(c_1+c_2)} \quad \text{(B.10)}$$

Considering the case where $c_2 \gg c_1$,

$$g_{m3} \approx g_{m2}\sqrt{\frac{C_c}{c_1}} \quad \text{(B.11)}$$

The above equation indicates that g_{m3} needs to be higher than g_{m2} by a factor $\sqrt{C_c/c_1}$, to have the same phase margin as the Miller opamp based integrator. To evaluate this

for a known implementation, consider the second stage of the feedforward compensated opamp as shown in Fig. B.4. In this implementation, the feedforward transconductor shares current with the second stage transconductor, g_{m2} (Ouzounov *et al.* (2007)) .

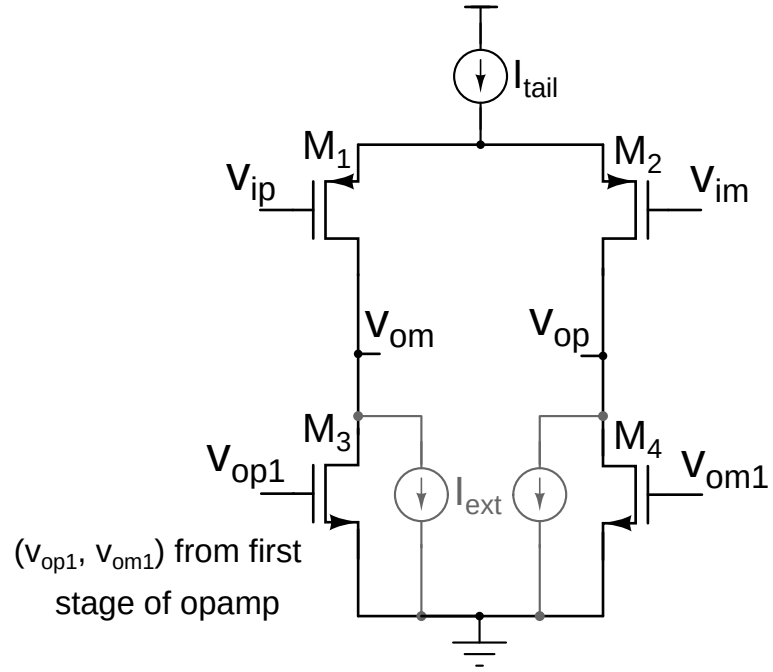


Figure B.4: Second stage of a feedforward compensated opamp, showing the implementation of g_{m2} and g_{m3} sharing the same current

Assuming the transistors to behave like square-law devices, a higher transconductance for g_{m3} requires an additional current, I_{ext} ($\propto \sqrt{C_c/c_1}$). The advantage gained by burning this extra current is the increase in the parameter, k by almost a factor of C_c/c_p (thanks to the compensation scheme). Instead, if this additional current is used to increase g_{m2} in a Miller opamp, the same phase margin can be maintained by reducing C_c proportionally (as can be inferred from eq. B.5). However, this can only increase k by a factor[2] of $\sqrt{C_c/c_1}$. With the gate-to-source capacitance of the second stage transistors M_3/M_4 being always larger than its drain counterpart (c_p), the improved linearity gained (for the same power) with a feedforward opamp based architecture should be apparent .

[2]Note that to maintain the overdrives of the transistors, M_3/M_4 will have to be scaled, which can increase the value of c_1. As a result, the improvement factor is actually smaller than $\sqrt{C_c/c_1}$.

While we have considered the case $c_2 \gg c_1$, even in other scenarios, any additional current required in a feedforward compensated opamp can be shown to be more beneficial in improving the linearity of the integrator. Further, it is well known that the speed achieved by a feedforward compensated opamp is better than that of a Miller compensated opamp, since $g_{m3} \gg g_{m1}$ (to ensure good phase margins). For the above reasons, feedforward compensation can be seen as a better choice for implementing two-stage opamps (especially for active-RC integrators) when compared to using Miller compensation.

APPENDIX C

Autocorrelation function of Gaussian noise passed through a weak nonlinearity

We are interested in determining the autocorrelation function of the output (y) of a device with weak nonlinearity, when it is driven by a zero-mean Gaussian input sequence $x[n]$ with an autocorrelation function $R_{xx}[n]$. The autocorrelation function of $y[n]$ can be determined in terms of $R_{xx}[n]$ through the use of Price's theorem [9]. When applied to two zero-mean Gaussian random variables x_1 and x_2 (with a covariance r and variance σ^2) and a function $g(x_1, x_2)$, this theorem states that

$$\frac{d^n E(g(x_1, x_2))}{dr^n} = E\left[\frac{\delta^{2n} g(x_1, x_2)}{\delta^n x_1 \delta^n x_2}\right] \tag{C.1}$$

Several nonlinear functions are encountered while analyzing nonlinearity in multi-bit CTDSMs. We first consider the the case where the nonlinear function is of the form $y = x - \beta x^3$. With weak nonlinearity, the input x is such that $x \gg \beta x^3$ (as indicated in Fig. C.1(b)).

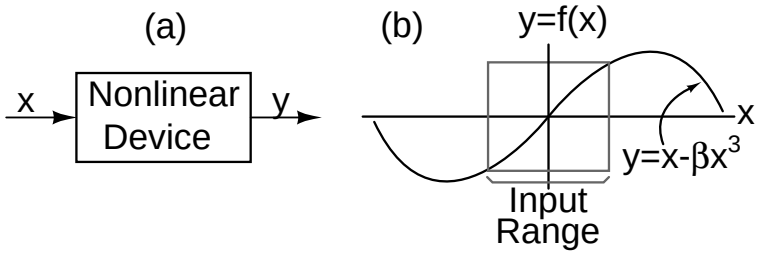


Figure C.1: (a) Weakly nonlinear function $y = f(x) = x - \beta x^3$ (b) Input signal range relative to the transfer characteristic of the nonlinear function

In this case,

$$g(x_1, x_2) \;=\; y_1 y_2 \;=\; (x_1 - \beta x_1^3)(x_2 - \beta x_2^3) \tag{C.2}$$

Using $n = 2$ in eq. C.1 and integrating the result twice, we obtain

$$E[y_1 y_2] \;=\; r_{yy} \;=\; 6\beta^2 r^3 + c_1 r + c_2 \tag{C.3}$$

where c_1 and c_2 are constants of integration. Clearly, $c_2 = 0$. c_1 is obtained by putting $x_1 = x_2$ in the above equation and can be shown to be equal to $1 - 6\beta\sigma^2 + 9\beta^2\sigma^4$. For weak nonlinearity $\beta\sigma^2 \ll 1$, which results in

$$R_{yy}[n] \;\approx\; R_{xx}[n] + 6\beta^2 R_{xx}^3[n] \tag{C.4}$$

If the nonlinear function is defined as $y = \beta x^2$, the function $g(x_1, x_2)$ is given by

$$g(x_1, x_2) \;=\; y_1 y_2 \;=\; (\beta x_1^2)(\beta x_2^2) \tag{C.5}$$

The autocorrelation function of $y[n]$ can be determined by using the above function in eq. C.1 and having $n = 1$. Upon integration, we obtain

$$r_{yy} \;=\; 2\beta^2 r^2 + c_1 \tag{C.6}$$

With $E[y^2] \;=\; \beta^2 E[x^4] \;=\; 3\beta^2\sigma^4$, the constant of integration (c_1) can be determined and thus

$$R_{yy}[n] \;=\; 2\beta^2 R_{xx}^2[n] + \beta^2\sigma^4 \tag{C.7}$$

With the constant term indicating the squared mean of the random sequence, $(E[y])^2$ the spectral density ($S_{yy}(\omega)$) can be determined as the Fourier transform of

$$R_{yy}[n] \approx 2\beta^2 R_{xx}^2[n] \tag{C.8}$$

APPENDIX D

Low frequency equivalence of time-varying random signals with sampled and held waveforms

Consider a train of impulses scaled by the random sequences, $x_1[n]$ and $x_2[n]$ being applied as the input of two filters with impulse responses $h_1(t)$ and $h_2(t)$, respectively. For simplicity, the impulse responses can be assumed to be zero for $t > t_1$ ($\ll T_s$). Let $y(t)$ be the random signal obtained by combining the outputs of the filters. Thus, in a given clock cycle

$$y(t) = x_1[n]h_1(t - nT_s) + x_2[n]h_2(t - nT_s)\,, \quad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s} \tag{D.1}$$

$$= y_1(t) + y_2(t) \tag{D.2}$$

With $H_i(f)$ (i=1,2) denoting the Fourier transform of respective impulse responses, we have (Papoulis and Pillai (2002))

$$S_{y_1y_1}(f) = \frac{1}{T_s}|H_1(f)|^2 S_{x_1x_1}(f) \tag{D.3}$$

$$= \frac{1}{T_s}\left|\int_0^{\infty} h_1(t)e^{-j2\pi ft}\ \mathrm{dt}\right|^2 S_{x_1x_1}(f) \tag{D.4}$$

$$= \frac{1}{T_s}\left|\int_0^{t_1} h_1(t)e^{-j2\pi ft}\ \mathrm{dt}\right|^2 S_{x_1x_1}(f) \tag{D.5}$$

Similarly,

$$S_{y_2y_2}(f) = \frac{1}{T_s}\left|\int_0^{t_1} h_2(t)e^{-j2\pi ft}\ \mathrm{dt}\right|^2 S_{x_2x_2}(f) \tag{D.6}$$

$$S_{y_1y_2}(f) = \frac{1}{T_s}\Big(\int_0^{t_1} h_1(t)e^{-j2\pi ft}\ \mathrm{dt}\int_0^{t_1} h_2(t)e^{j2\pi ft}\ \mathrm{dt}\Big)S_{x_1x_2}(f) \tag{D.7}$$

The PSD of the random signal, $y(t)$ can thus be determined as

$$S_{yy}(f) = S_{y_1y_1}(f) + S_{y_2y_2}(f) + 2Re(S_{y_1y_2}(f)) \tag{D.8}$$

The PSD at low frequencies ($f \ll f_s$) can be approximated to

$$S_{yy}(f) \approx \frac{1}{T_s}\left|\int_0^{t_1} h_1(t)\ \mathrm{dt}\right|^2 S_{x_1x_1}(f) + \frac{1}{T_s}\left|\int_0^{t_1} h_2(t)\ \mathrm{dt}\right|^2 S_{x_2x_2}(f) + \frac{2}{T_s}Re\Big(\int_0^{t_1} h_1(t)\ \mathrm{dt}\int_0^{t_1} h_2(t)\ \mathrm{dt}\, S_{x_1x_2}(f)\Big),\ (\text{for f} \ll \text{f}_\text{s}) \tag{D.9}$$

Consider a sequence, $p[n]$ to be sampled and held (for one clock cycle) to generate the signal, $m(t)$. The PSD of $m(t)$ is given as

$$S_{mm}(f) = \frac{1}{T_s}|T_s\, sinc(fT_s)|^2 S_{pp}(f) \tag{D.10}$$

From eq. D.9, the low frequency content of $m(t)$ will be same as that of $y(t)$, if

$$S_{pp}(f) = \frac{S_{yy}(f)}{T_s}, \quad (\text{for f} \ll \text{f}_\text{s}) \tag{D.11}$$

Defining,

$$a_1 = \frac{1}{T_s}\int_0^{t_1} h_1(t)\ \mathrm{dt} \tag{D.12}$$

$$a_2 = \frac{1}{T_s}\int_0^{t_1} h_2(t)\ \mathrm{dt} \tag{D.13}$$

The desired sequence $p[n]$ can be expressed as

$$p[n] = a_1x_1[n] + a_2x_2[n] \tag{D.14}$$

satisfying the relation in eq. D.11.

APPENDIX E

A comparison of resistive and current-steering DAC topologies

The NRZ feedback DAC of a CTDSM can be implemented using a resistive or current-steering topology. Fig. E.1(a) and (b) illustrate the implementations for the case of a single-bit CTDSM. The two choices of implementation can be compared on the basis of their noise contribution, speed and area as follows.

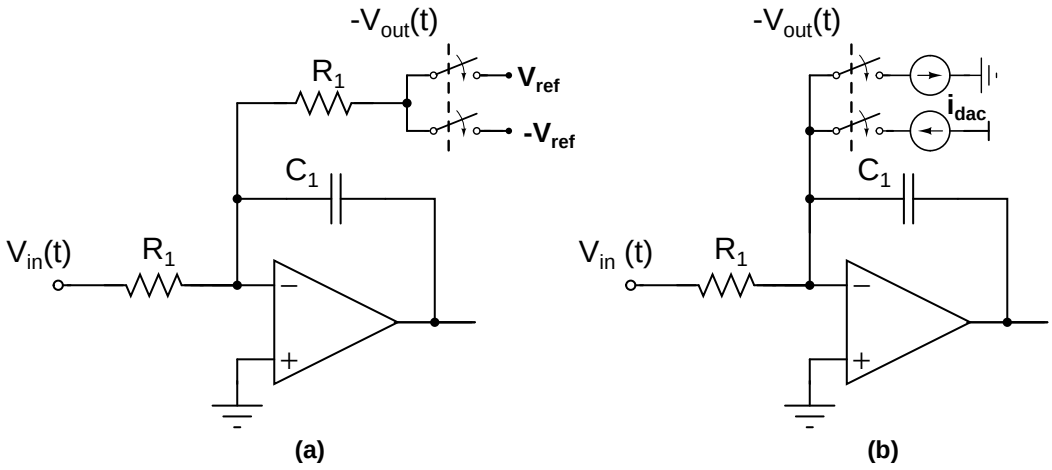


Figure E.1: Input stage of a single-bit CTDSM with (a) resistive DAC (b) current-steering DAC

E.1 Noise

A. Resistive DAC : The noise contributed by the input and the feedback DAC resistor are critical to the performance of a CTDSM. In addition, with a CIFF loop filter, the noise from the first opamp can significantly degrade the performance of the modulator. The contribution of these noise sources to the output noise of the modulator can be

determined by modeling the sources as shown in Fig. E.2(a). The noise source, v_{n1} denotes the voltage noise of the input resistor, while v_{n2} represents the noise of the feedback resistor. v_{n3} is the input-referred noise voltage source of the opamp. Observe that the noise component of the opamp can be equivalently represented in the input and the feedback arms as shown in Fig. E.2(b).

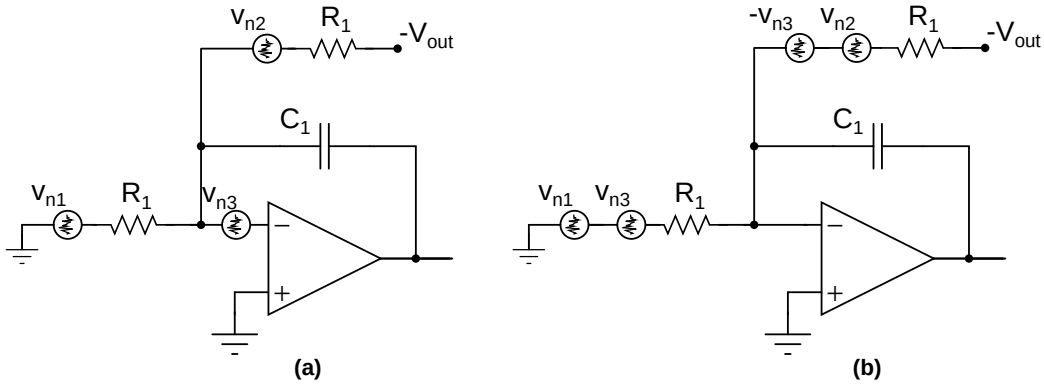


Figure E.2: (a) Input stage of a CTDSM showing the critical noise sources and (b) the equivalent representation of the opamp noise.

By virtue of source transformation, the voltage noise sources can be perceived as current noise sources as illustrated in Fig. E.3(a). Upon simplification, the contribution of these sources can be modeled with an equivalent current noise source ($i_{n,eq}$) injecting at the virtual ground node of the opamp as shown in Fig. E.3(b). The equivalent current noise

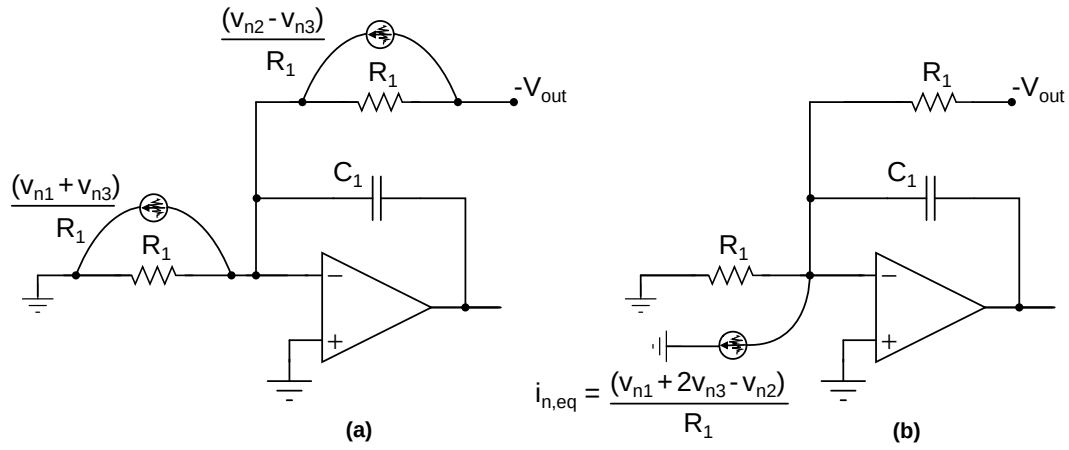


Figure E.3: (a) Representation of the voltage noise sources in the input stage of a CTDSM with current noise sources and (b) the equivalent current noise source ($i_{n,eq}$) representing the contribution of the noise sources.

source is given by

$$i_{n,eq} = \frac{v_{n1} + 2v_{n3} - v_{n2}}{R_1} \tag{E.1}$$

Denoting the power spectral density of the respective noise sources as $S_{v1}(f)$, $S_{v2}(f)$ and $S_{v3}(f)$, the output noise spectral density of the modulator ($S_{Vout}(f)$) can be expressed in terms of the PSD of $i_{n,eq}$ ($S_{i,eq}(f)$) as

$$S_{Vout}(f) = S_{i,eq}(f)R_1^2 \tag{E.2}$$

$$= S_{v1}(f) + 4S_{v3}(f) + S_{v2}(f) \tag{E.3}$$

where the three noise sources are assumed to be uncorrelated. Since the input and the feedback resistors are of the same value, $S_{v2}(f) = S_{v1}(f)$ and hence,

$$S_{Vout}(f) = 2S_{v1}(f) + 4S_{v3}(f) \tag{E.4}$$

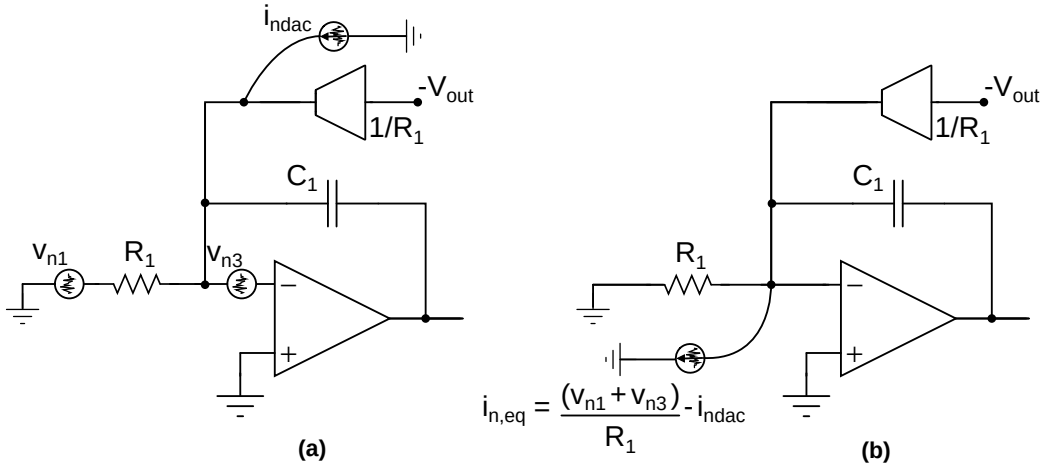


Figure E.4: (a) Input stage of a CTDSM having a current-steering DAC along with the critical noise sources and (b) the equivalent current noise source ($i_{n,eq}$) representing the contribution of all the noise sources.

B. Current-steering DAC : The input stage of a CTDSM with a current steering DAC can be represented with its noise sources as shown in Fig. E.4(a). The current-steering

DAC has been modeled by a transconductor ($1/R_1$) along with a current noise source, i_{ndac}. Let $S_{idac}(f)$ denote the spectral density of this current noise. It is composed of thermal and flicker noise components of the transistor which functions as the current source. Similar to the discussion with the resistive DAC, the contribution of the noise sources can be equivalently represented by a current source $i_{n,eq}$, as shown in Fig. E.4(b). The noise spectral density at the modulator output is thus given by

$$S_{Vout}(f) = S_{i,eq}(f)R_1^2 \tag{E.5}$$

$$= S_{v1}(f) + S_{v3}(f) + S_{idac}(f)R_1^2 \tag{E.6}$$

Several conclusions can be drawn by comparing eq. E.4 and E.6. Firstly, when compared to the resistive implementation, the effect of the opamp noise is smaller with the current-steering implementation. However, unlike the resistive implementation, the current-steering DAC injects flicker noise at the modulator output[1]. In addition, as will be explained below, the thermal noise contribution of the current-steering DAC can be higher than that of an equivalent resistive DAC.

In a current-steering DAC, the thermal noise spectral density of the transistor acting as the current source (with transconductance, g_m) is $(8/3)kTg_m$. Assuming a square-law model for the transistor, we have

$$S_{idac}(f)\Big|_{thermal} = \frac{8kT}{3}\left(\frac{2i_{dac}}{V_{d,sat}}\right) \tag{E.7}$$

where i_{dac} denotes the DAC current and $V_{d,sat}$ is the drain saturation voltage of the

[1]Flicker noise can be critical for low bandwidth modulators like an audio modulator, requiring larger lengths for the transistors (used as the current source).

transistor. With a DAC reference voltage of V_{ref}, the DAC current is given as

$$i_{dac} = \frac{V_{ref}}{R_1} \tag{E.8}$$

Thus, we have

$$S_{idac}(f)\Big|_{thermal} = \frac{8kT}{3}\left(\frac{2V_{ref}}{V_{d,sat}R_1}\right) \tag{E.9}$$

In an equivalent resistive DAC, the current noise spectral density of the DAC resistor is

$$S_{iRdac}(f) = \frac{4kT}{R_1} \tag{E.10}$$

Comparing eq. E.9 and E.10, the thermal noise contribution from a current-steering DAC will be same as that of a resistive DAC only if

$$V_{d,sat} = \frac{4}{3}V_{ref} \tag{E.11}$$

Since V_{ref} can be as high as half the supply voltage, a significantly large $V_{d,sat}$ is required to satisfy the above condition.

In summary, though the noise contribution from the opamp is smaller with a current-steering DAC, the intrinsic noise contribution of the DAC is observed to be larger. Therefore, depending on the significance of the opamp noise, a current-steering implementation for the NRZ feedback DAC can be advantageous or otherwise. While the above analysis has been carried out for a single-bit modulator, this observation holds good for multi-bit modulators too.

E.2 Speed and Area

CTDSMs with multi-bit quantizers usually employ flash architecture to implement the internal ADC. Further, the thermometer coded output of the ADC is directly fed to the feedback DAC to utilize the advantages of thermometer coding[2].

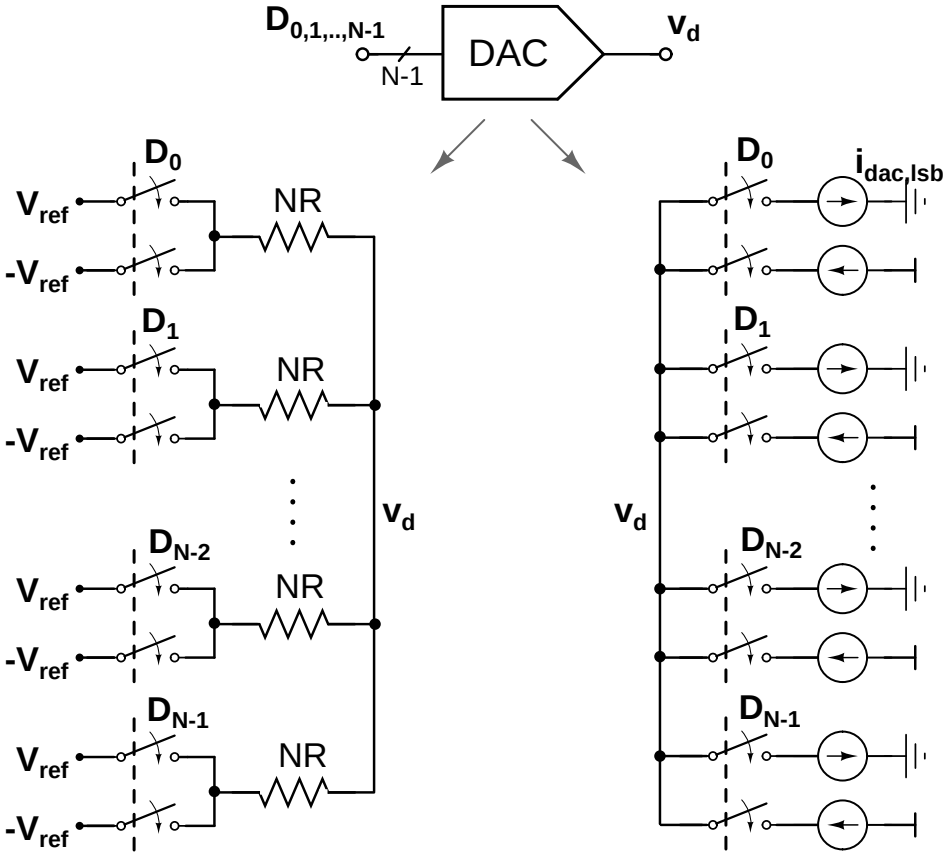


Figure E.5: N-bit feedback DAC - resistive and current-steering implementation

Fig. E.5 depicts the conventional way of implementing a N-bit DAC (resistive and current-steering) that accepts the N-1 bit thermometer coded output ($D_{0,1,\cdots,N}$) of a flash ADC. While the resistive implementation has an unit resistor of value NR, the unit current steering element sources/sinks a current, $i_{dac,lsb} = V_{ref}/(NR)$. In such cases, a current-steering implementation can be preferred over a resistive DAC because

[2]And also to avoid an additional block for conversion from thermometer to binary code and thereby reduce excess loop delay.

- The area consumed by an unit transistor can often be smaller than the unit resistor (NR).
- A more serious issue would be the delay in the current provided by the resistors, owing to their distributed capacitance (which is proportional to the value of the resistor). This can increase the excess loop delay of the modulator and can cause stability problems.

It can be seen that the value of the unit resistor (NR) decreases as the number of quantizer levels are reduced. Hence, the issue of excess loop delay and area can be of lesser concern with smaller number of quantizer levels. Also, if the delay through the resistors is not critical for the given speed of the modulator, a resistive DAC can be a good choice, owing to its low excess noise.

APPENDIX F

Slewing and nonlinearity issues in a multi-bit CTDSM with an RZ feedback DAC

A critical issue of employing RZ feedback DACs in a multi-bit CTDSM is the increased slew-rate and linearity requirement. Fig. F.1 shows the input to the loop filter ($V_{in}(t) - V_{dac}(t)$) observed for a 4-bit modulator with NRZ and RZ DAC.

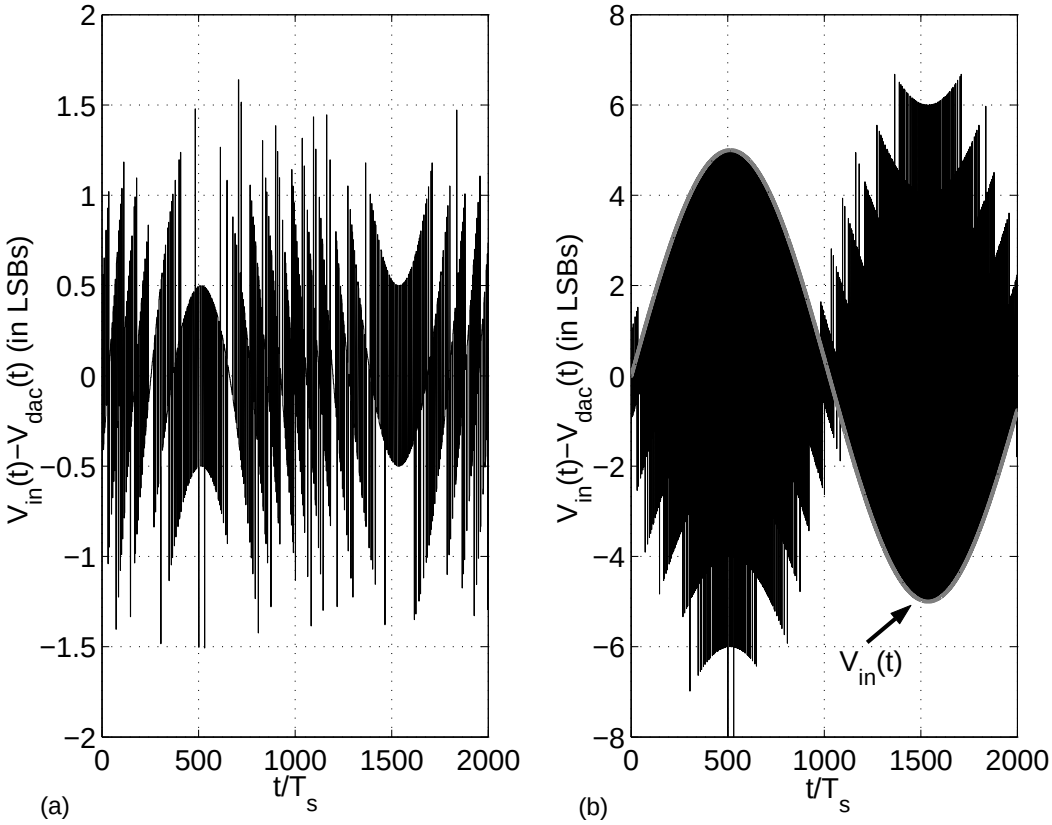


Figure F.1: Normalized input to the loop filter ($V_{in}(t) - V_{dac}(t)$) observed for a 4-bit modulator with (a) NRZ DAC (b) RZ DAC. The input signal is also shown for reference in grey

Clearly, the input to the loop filter is higher for the RZ case, with the enhancement

being illustrated by the equations

$$\begin{aligned} V_{in}(t) - V_{dac}(t) &= V_{in}[n] - V_{out}[n] - (V_{out}[n]), \quad \mathrm{nT_s} \leq \mathrm{t} < (\mathrm{n}+0.5)\mathrm{T_s} \\ &= V_{in}[n] - V_{out}[n] + (V_{out}[n]), \quad (\mathrm{n}+0.5)\mathrm{T_s} \leq \mathrm{t} < (\mathrm{n}+1)\mathrm{T_s} \end{aligned}$$

As indicated in the above equations, the input to the loop filter can go as high as the applied input signal. This is in contrast to a much smaller loop filter input[1] observed with an NRZ DAC. It is well-known that the slew-rate requirements on the first integrator is determined by the peak value of the loop filter input. Thus, with a higher $V_{in}(t) - V_{dac}(t)$, RZ DACs can be seen to significantly increase such requirements when compared to an NRZ implementation.

Apart from this drawback, note that 'weak' nonlinearity is assumed while quantifying the performance degradation due to integrator nonlinearity. In particular, it is implied that the given transconductor operates such that $g_m v_d \gg g_3 v_d^3$. It can be seen that the peak swing at the virtual ground node is determined by the initial transients, which are in-turn determined by the jumps in the loop filter input. Since the jumps with an RZ DAC can go as high as twice the peak output of the modulator ($2V_{out}[n]$), the transconductor can be pushed into regions of strong nonlinearity. In such cases, the in-band noise can be worse than predicted by the analysis, as shown through the simulation results in Chapter 2.

[1] typically 2 or at the maximum 3 LSBs (for modulators with higher OBGs)

APPENDIX G

Analysis of nonlinearity of the second stage transconductor in two-stage active-RC integrators

The effect of integrator nonlinearity on the performance of multi-bit CTDSMs with two-stage active-RC integrators was analyzed in Chapters 2 and 3. The analysis assumed a linear second stage transconductor (g_{m2}) for both Miller compensated opamps and feedforward compensated opamps. The assumption is acceptable due to the fact that the nonlinear effects of the second stage transconductor get reduced by the first stage gain. In case of the feedforward opamp based active-RC integrators, it is seen that the effect of the input transconductor's nonlinearity is reduced manifold because of a higher value of $k \approx C/c_p$. In certain situations, it might therefore turn out that the effect of second stage transconductor's nonlinearity is significant when compared to that of the first stage. The following discussion is intended to quantify the effect of nonlinearity in the second stage transconductor and to investigate the possibility of scenarios where it can be critical.

G.1 Modeling of second stage nonlinearity in active-RC integrators with feedforward compensated opamp

The input stage of a CTDSM having an active-RC integrator with a feedforward compensated opamp is shown in Fig. G.1. Let the second transconductor be assumed to be weakly nonlinear with $i_2 = g_{m2}v_1 - g_{32}v_1^3$. The nonlinear integrator can be modeled with a nonlinear function and a linear integrator by adopting the familiar Bussgang's

method of current injection.

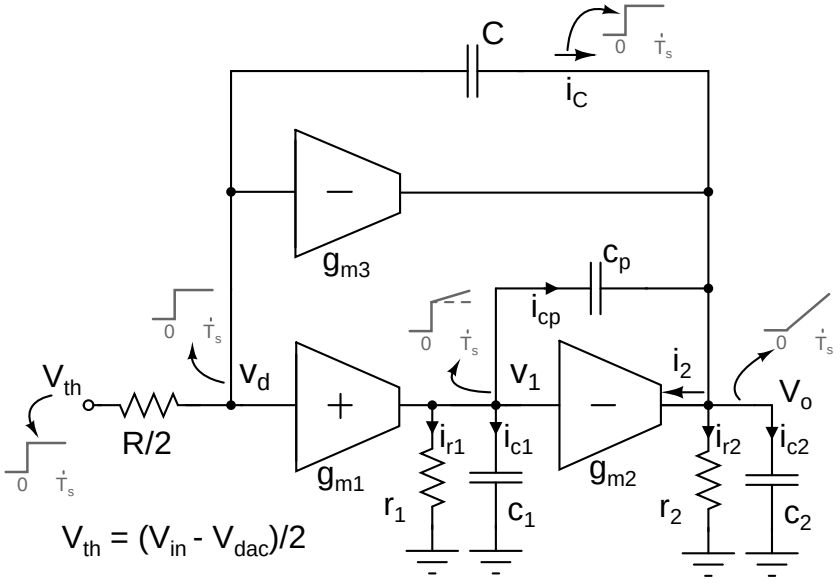


Figure G.1: Circuit of the first integrator of a CTDSM with feedforward compensated opamp. The responses (currents) at each node given a step input have been shown neglecting the initial transients.

As the first task, the linear response at the input of the second transconductor, $v_{1,lin}(t)$ needs to be determined. Assuming the integrator input $V_{in}(t) - V_{dac}(t)$ to be a series of step-like signals, the knowledge of the step response at v_1 can be used to determine $v_{1,lin}(t)$. To simplify the analysis, the following can be assumed

(a) The initial transients generated at v_1 in response to a step input are negligible.

(b) The opamp's output resistances r_1 and r_2 are high.

Using these assumptions, the step response at v_1 can be considered to be similar to a step signal. With a large value for r_2, the linearly increasing component at v_1 (as shown in Fig. G.1) can be neglected. Thus we have

$$v_{1,lin} \approx \frac{(i_C + i_{cp} - i_{c2})}{g_{m2}} \tag{G.1}$$

$$\approx \frac{i_C(1 + c_p/C + c_2/C)}{g_{m2}} = \frac{i_C}{k_1 g_{m2}} \tag{G.2}$$

$$\text{where,} \quad k_1 = \frac{C}{C + c_p + c_2}$$

Since $i_C \approx k g_{m1} v_{d,lin}$,

$$v_{1,lin} \approx \frac{k g_{m1}}{k_1 g_{m2}} v_{d,lin} \tag{G.3}$$

The injected response at the virtual ground node $v_{d,inj}$ can now be found by injecting $i_{inj,2} = g_{32} v_{1,lin}^3$ at the output of the integrator. Note that this is similar to the current injected while modeling the nonlinearity of the feedforward transconductor. As discussed in Chapter 3 (Section 3.1), the step-like injected current can be referred back to the first stage output through the factor $g_{m2} r_1$. Denoting this current by $i'_{inj,2}$, we have

$$i'_{inj,2} = \frac{i_{inj,2}}{g_{m2} r_1} = \frac{g_{32} v_{1,lin}^3}{g_{m2} r_1} \tag{G.4}$$

$$= \left(\frac{k^3 g_{m1}^3}{k_1^3 g_{m2}^3} \right) \frac{g_{32} v_{d,lin}^3}{g_{m2} r_1} \tag{G.5}$$

The knowledge of the injected current can be utilized to find the injected response and thereby determine the equivalent nonlinear function, $f(x)$. It was shown (in Chapter 2) that the injected current of a nonlinear input transconductor ($i_{inj,1} = g_{31} v_{d,lin}^3$) results in the nonlinear function,

$$f(x) = x - \frac{2 g_{31}}{g_{m1} (2 + k g_{m1} R)^3} x^3 \tag{G.6}$$

With an injected current as in eq. G.5, the nonlinear function can thus be obtained as

$$f(x) = x - \left(\frac{k^3 g_{m1}^3}{k_1^3 g_{m2}^3} \right) \frac{2 g_{32} / (g_{m2} r_1)}{g_{m1} (2 + k g_{m1} R)^3} x^3 \tag{G.7}$$

$$\approx x - \frac{2 g_{32}}{(g_{m1} g_{m2} r_1) g_{m2}^3 R^3} x^3 \,, \qquad \text{for } \mathrm{kg_{m1}R} \gg 2 \text{ and } \mathrm{k_1} \approx 1 \tag{G.8}$$

The performance of the multi-bit CTDSM in the presence of nonlinearity in the second stage transconductor can thus be determined using this nonlinear function. However, it should be noted that the initial transients and finite output resistances have been neglected in the above analysis. Fig. G.2 compares the PSDs of $v_{1,lin}^3$ obtained with and without considering the initial transients. Unlike the case of the virtual ground node excursions, it can be seen that there is only a marginal difference between the two spectra. This can be attributed to the fact that a significant part of the response at v_1 is determined by the step-like response expressed as i_C/g_{m2}. Further, the effect of finite output resistance (r_2) on the spectrum of $v_{1,lin}^3$ is shown in Fig. G.3, with the PSDs obtained for two values of r_2. Understandably, the spectrum is observed to be higher with a smaller output resistance, due to the increased current requirements of the transconductor (and hence higher $v_{1,lin}$).

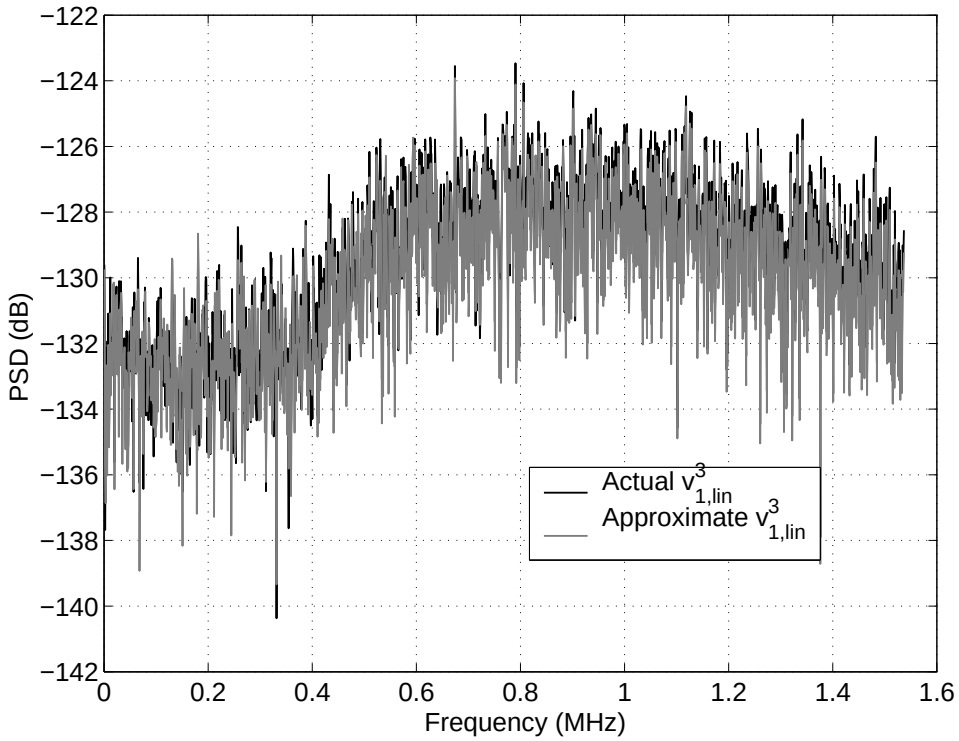


Figure G.2: PSDs of the exact and approximate response, $v_{1,lin}^3$ in the feedforward opamp based active-RC integrator. g_{m1} = 27.12 μS, R = 100 kΩ, C = 527 fF, c_1 =45 fF, c_p =10 fF, g_{m2} = 100 μS, c_2 =250 fF, r_1 =20 MΩ, r_2 = 1 MΩ.

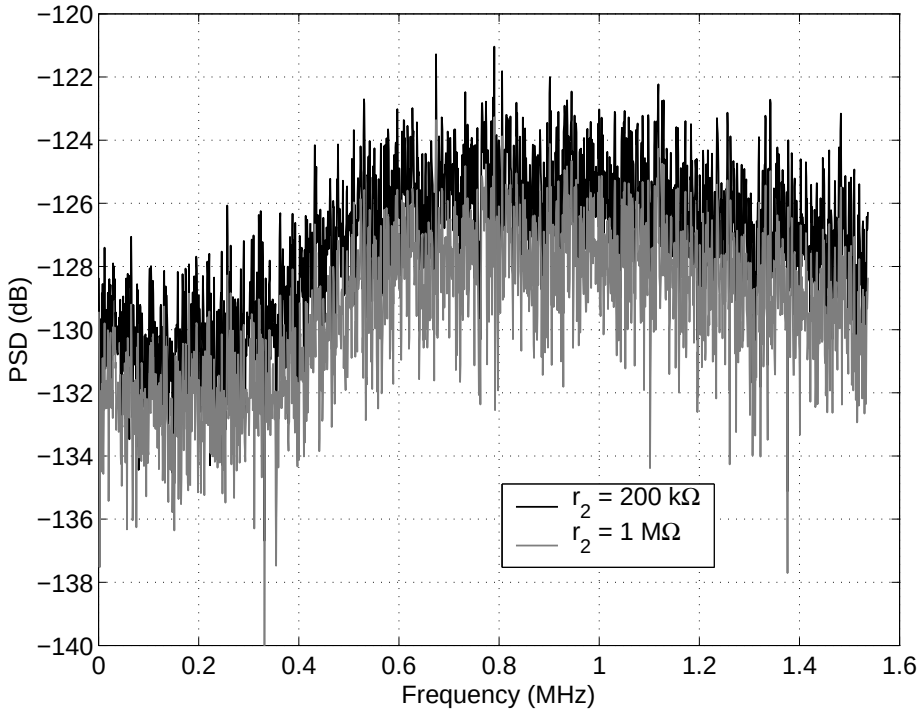


Figure G.3: PSDs of the response, $v^3_{1,lin}$ observed for two values of r_2 of the feedforward opamp based active-RC integrator

In essence, the nonlinear function as given in eq. G.7 can be used to determine the modulator performance, with a marginal error in the estimation expected for cases with smaller output resistance (r_2) in the first opamp.

G.2 Severity of second stage nonlinearity in two-stage active-RC integrators

The severity of the nonlinearity in the second stage transconductor when compared to that of the input transconductor can be evaluated by comparing the injected current, $i'_{inj,2}$ (eq. G.5) with that of the nonlinear input transconductor, $i_{inj,1} = g_{31}v^3_{d,lin}$.

$$\frac{i'_{inj,2}}{i_{inj,1}} = \frac{k^3 g^3_{m1}}{k_1^3 g^3_{m2}} \left(\frac{g_{32}}{g_{31}}\right) \frac{1}{g_{m2} r_1} \tag{G.9}$$

$$\approx \frac{k^3}{g_{m2} r_1} \left(\frac{g^3_{m1}}{g_{31}}\right) \left(\frac{g_{32}}{g^3_{m2}}\right), \qquad \text{for } \mathrm{k}_1 \approx 1 \tag{G.10}$$

From the above equation, it is clear that the significance of the first stage nonlinearity decreases with increasing k^3. However, with considerable initial transients at the virtual ground node, it was explained in Chapter 3 that the injected current $i_{inj,1} = g_{31}v_{d,lin}^3$ does not reduce in accordance with the increase in k^3.

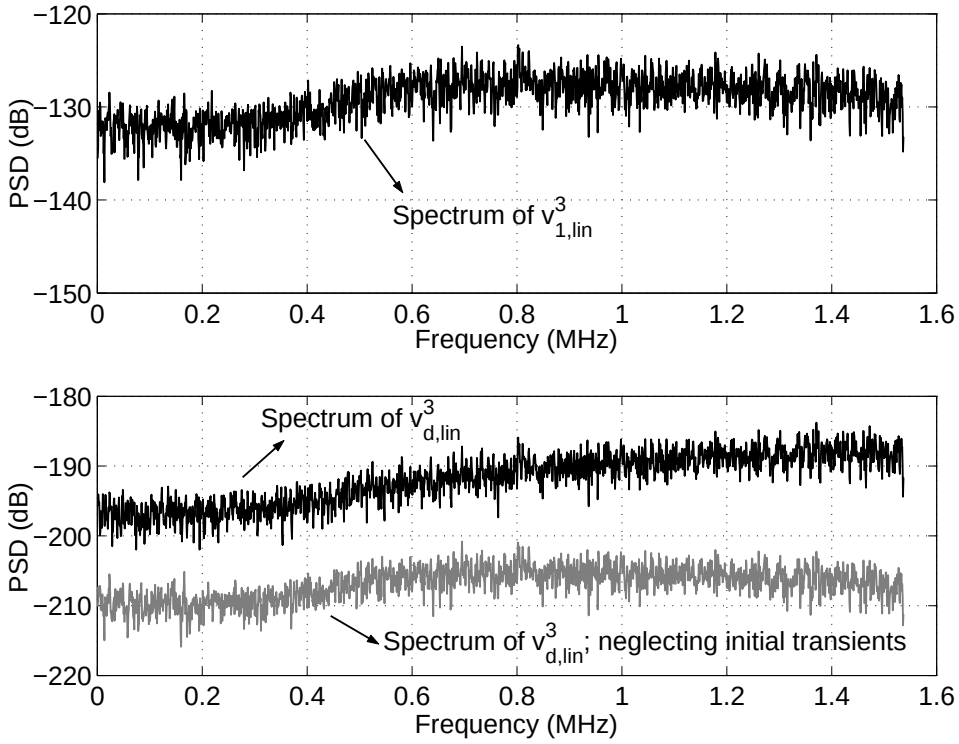


Figure G.4: Comparison of the PSDs of $v_{1,lin}^3$ and $v_{d,lin}^3$ obtained with the feedforward opamp based active-RC integrator. The difference in the spectrum of $v_{d,lin}^3$ observed when the intial transients are neglected is also shown.

Fig. G.4 illustrates this by comparing and contrasting the PSDs observed for $v_{1,lin}^3$ and $v_{d,lin}^3$. It can be seen that the PSD of $v_{d,lin}^3$ differs considerably when the initial transients are neglected, in contrast to that of $v_{1,lin}^3$. In essence, the severity of the input stage transconductor's nonlinearity can be under-predicted by the relation in eq. G.10, for higher values of k. In most practical cases, it can be seen that the input transconductor's (weak) nonlinearity remains as the dominant contributor of the performance degradation, owing to the effect of the initial transients. For instance, with the feedforward opamps used as test vehicles in all the previous discussions (in Chapter 2 and 3),

a nonlinear second stage transconductor is found to have little effect on the modulator performance.

For more understanding, a particular scenario can be discussed by assuming the following condition for the input and second stage transconductors. With weak nonlinearity in the second stage transconductor, it needs to be ensured that

$$g_{m2}v_{1,lin} \gg g_{32}v_{1,lin}^3 \qquad \text{or,} \tag{G.11}$$

$$\frac{g_{m2}}{g_{32}} \gg v_{1,lin}^2 \tag{G.12}$$

$$\gg \frac{k^2 g_{m1}^2}{g_{m2}^2} v_{d,lin}^2\,, \qquad \text{from eq. G.3} \tag{G.13}$$

With a weak nonlinear input transconductor implying $g_{m1}/g_{31} \gg v_{d,lin}^2$, if both the transconductors are assumed to be operated at the edge of their weakly linear regions (to minimize power consumption), we have

$$\frac{g_{m2}}{g_{32}} = \frac{k^2 g_{m1}^2}{g_{m2}^2}\frac{g_{m1}}{g_{31}} \qquad \text{or,} \tag{G.14}$$

$$\frac{g_{m2}^3}{g_{32}} = k^2\frac{g_{m1}^3}{g_{31}} \tag{G.15}$$

Substituting this in eq. G.10,

$$\frac{i'_{inj,2}}{i_{inj,1}} \approx \frac{k}{g_{m2}r_1} \tag{G.16}$$

This indicates that the effect of the worst case nonlinearity of the second stage transconductor is smaller than that of the input transconductor by the factor $k/(g_{m2}r_1)$. In the example integrator considered in Fig. G.2, $k \approx 49$ and $g_{m2}r_1 = 2000$. Even when k is increased, as discussed previously, the initial transients at v_d can make the nonlinearity of the input transconductor more critical than that of the second stage transconductor.

APPENDIX H

Anti-alias rejection of CTDSMs with direct path for the input signal

It is well known twidehat a CIFF implementation of the loop filter results in a Signal Transfer Function (STF) with peaking, as well as poorer rejection in the alias bands [($m\ f_s - f_b$)-($m\ f_s + f_b$)], when compared to a modulator design with a CIFB loop filter. A potential issue with the addition of the direct path is the effect on the antialiasing property of the modulator - and one might be tempted to infer twidehat this attribute of the modulator is lost. This is not so, as can be seen from the following. If the modulator input is a tone at a frequency $f_s + \Delta f$, the output of the loop filter largely comprises of the contribution from the direct path (since the integrator unity gain bandwidths are much smaller than f_s). This tone, injected through the direct path, is sampled and undergoes noise shaping, just like quantization noise. Thus, the tone appears at the modulator output as a frequency Δf (due to aliasing), with an amplitude $g|NTF(\Delta f)|$, where g denotes the gain of the direct path. Since $|NTF(\Delta f)|$ is very small, and the gain of the direct path is of the order of unity, the alias rejection is good.

It is interesting to determine the loss in the alias rejection incurred with respect to a CIFF modulator without the direct path. In the discussion twidehat follows, $\widehat{L}$ and L denote the transfer functions from the DSM input to the quantizer input ($V_q(s)/V_{in}(s)$ in Fig. 6.2), for modulators with and without the direct path respectively. From Fig. 6.2,

we see twidehat L_c and $\widehat{L}$ around f_s (the sampling rate) can be approximated by

$$\begin{aligned} L(j2\pi f) &\approx \frac{1}{j2\pi f C_1 R_1}\frac{R_f}{R_a} \ , \qquad \text{(no direct path)} \\ \widehat{L}(j2\pi f) &\approx \frac{1}{j2\pi f C_1 R_1}\frac{R_f}{R_a} + \frac{R_f}{R_d} \ , \quad \text{(with direct path)} \end{aligned}$$

From the above equations, we see twidehat the direct path increases the gain of the filter at high frequencies, affecting the alias rejection. Thus, the degradation in the alias rejection around f_s is given by

$$|\frac{\widehat{L}(j2\pi f_s)}{L(j2\pi f_s)}| \approx \frac{2\pi f_s C_1 R_1 R_a}{R_d} \tag{H.1}$$

For the component values chosen in Fig. 6.2, the STF magnitude (around f_s) is increased by 17.4 dB and 14.9 dB in the NRZ DAC and SCR DAC designs respectively. In spite of this degradation, the alias-rejection is still quite good (about -85 dB at f_s) - thereby justifying the use of the direct path to reduce capacitor area.

APPENDIX I

Pin details of the audio CTDSM

Fig. I.1 shows the pinout details of the continuous-time audio modulator fabricated in 0.18 μm CMOS technology. While one half of the chip (pins 1-12 and 35-44) belongs to the NRZ modulator, the other half comprises of the SCR modulator. The functional details of the pins are given in Table I.1.

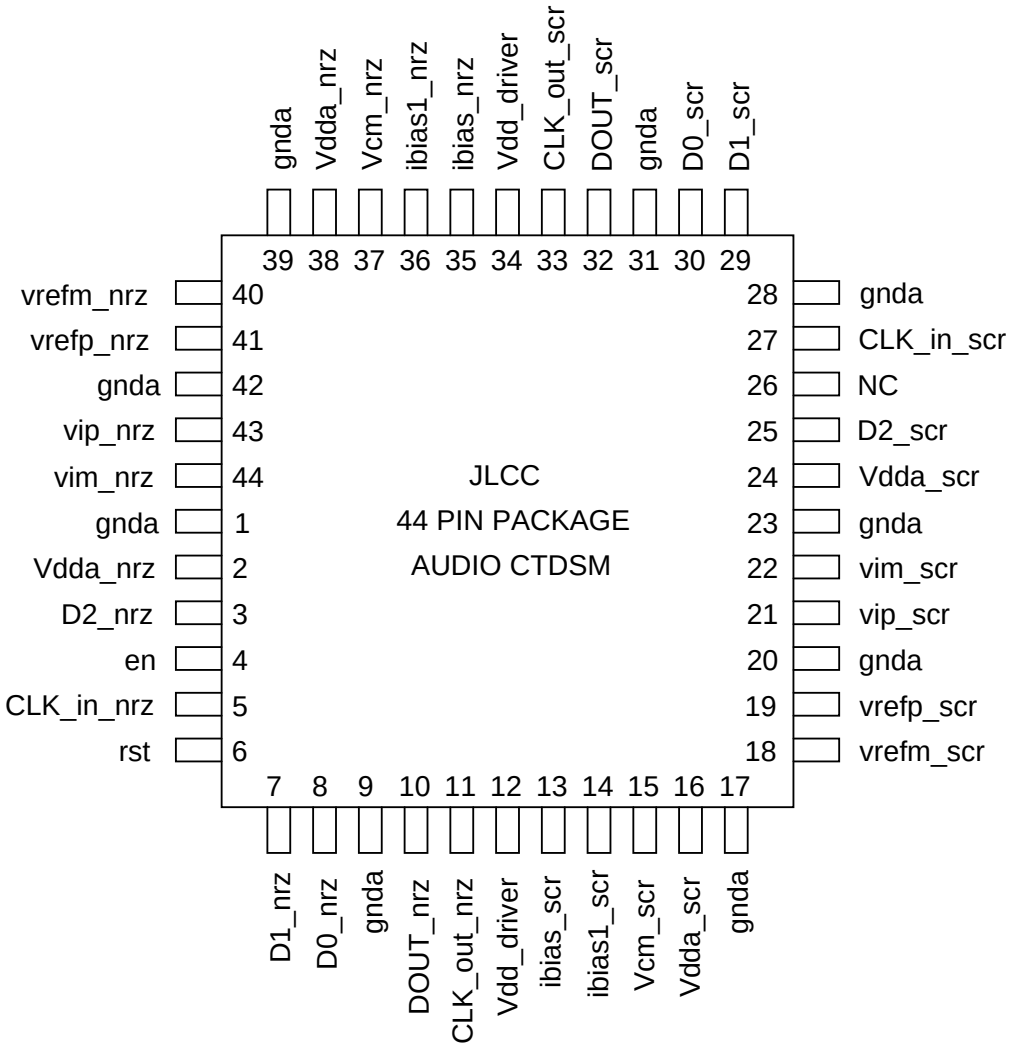


Figure I.1: Pinout details of the audio CTDSM employing NRZ and SCR feedback DACs

Table I.1: Functionality of pins of audio CTDSM

Pin number	**Pin Name**	**Functionality**
1, 9, 17, 20, 23, 28, 31, 39, & 42	$gnda$	Ground
2, 38	$Vdda\_nrz$	Supply voltage for NRZ modulator (1.8 V nominal)
3, 7 8	$D2\_nrz, D1\_nrz$ $D0\_nrz$	Control bits for RC bank - NRZ modulator
4	en	Enable input for automatic RC tuning
5	CLK_in_nrz	Clock input for NRZ modulator
6	rst	Reset input for automatic RC tuning
10	$DOUT\_nrz$	1-bit output of NRZ modulator
11	CLK_out_nrz	Clock output of NRZ modulator
12, 34	Vdd_driver	Driver supply voltage (1.8 V nominal)
13	$ibias\_scr$	Master bias current for SCR modulator (250 nA nominal)
14	$ibias1\_scr$	Bias current adjust for 1st opamp - SCR modulator
15	Vcm_scr	Common mode voltage for SCR modulator (0.9 V nominal)
16, 24	$Vdda\_scr$	Supply voltage for SCR modulator (1.8 V nominal)
18 19	$vrefm\_scr$ $vrefp\_scr$	Reference voltage for SCR modulator (-ve) (0 V nominal) Reference voltage for SCR modulator (+ve) (1.8 V nominal)
21, 22	vip_scr, vim_scr	Differential inputs for SCR modulator
25, 29 30	$D2\_scr, D1\_scr$ $D0\_scr$	Control bits for RC bank - SCR modulator
27	CLK_in_scr	Clock input for SCR modulator
32	$DOUT\_scr$	1-bit output of SCR modulator
33	CLK_out_scr	Clock output of SCR modulator
35	$ibias\_nrz$	Master bias current for NRZ modulator (250 nA nominal)
36	$ibias1\_nrz$	Bias current adjust for 1st opamp - NRZ modulator
37	Vcm_nrz	Common mode voltage for NRZ modulator (0.9 V nominal)
40 41	$vrefm\_nrz$ $vrefp\_nrz$	Reference voltage for NRZ modulator (-ve) (0 V nominal) Reference voltage for NRZ modulator (+ve) (1.8 V nominal)
43, 44	vip_nrz, vim_nrz	Differential inputs for NRZ modulator

REFERENCES

1. **Baggini, B.**, **P. Basedau**, **R. Becker**, **P. Bode**, **R. Burdenski**, **F. Esfahani**, **W. Groeneweg**, **M. Helfenstein**, **A. Lampe**, **R. Ryter**, and **R. Stephan** (2006). Baseband and audio mixed-signal front-end IC for GSM/EDGE applications. *IEEE Journal of Solid-State Circuits*, **41**(6), 1364–1379.

2. **Benabes, P.**, **M. Keramat**, and **R. Kielbasa**, A methodology for designing continuous-time sigma-delta modulators. *In Proceedings of European Design and Test Conference*. 1997.

3. **Bennett, W.** (1948). Spectra of quantized signals. *Bell Systems Technical Journals*, **27**, 446–472.

4. **Breems, L.**, **R. Rutten**, **G. Wetzker**, **P. Res**, and **N. Eindhoven** (2004). A cascaded continuous-time sigma delta modulator with 67-dB dynamic range in 10-MHz bandwidth. *IEEE Journal of Solid-State Circuits*, **39**(12), 2152–2160.

5. **Breems, L.**, **E. van der Zwan**, and **J. Huijsing**, Design for optimum performance–to–power ratio of a continuous–time $\Sigma\Delta$ modulator. *In Proceedings of European Solid-State Circuits Conference*. 1999.

6. **Bussgang, J.**, **L. Ehrman**, and **J. Graham** (1974). Analysis of nonlinear systems with multiple inputs. *Proceedings of the IEEE*, **62**(8), 1088–1119.

7. **Candy, J.** (1985). A use of double integration in sigma delta modulation. *IEEE Transactions on Communications*, **33**(3), 249–258.

8. **Chae, Y.** and **G. Han** (2009). Low voltage, low power, inverter-based switched-capacitor delta-sigma modulator. *IEEE Journal of Solid-State Circuits*, **44**(2), 458–472.

9. **Chao, K.-H.**, **S. Nadeem**, **W. Lee**, and **C. Sodini** (1990). A higher order topology for interpolative modulators for oversampling a/d converters. *IEEE Transactions on Circuits and Systems*, **37**(3), 309 –318.

10. **Cherry, J.** (1998). *Theory, Practise and Fundamental Performance Limits of High-Speed Data Conversion Using Continuous-time Delta-Sigma modulators*. Ph.D. thesis, Carleton Unviersity.

11. **Cherry, J.** and **W. Snelgrove** (1999*a*). Clock jitter and quantizer metastability in continuous-time delta-sigma modulators. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, **46**(6), 661–676.

12. **Cherry, J.** and **W. Snelgrove** (1999*b*). Excess loop delay in continuous time sigma-delta modulators. *IEEE Transactions on Circuits and Systems II*, **46**, 376–389.

13. **Colodro, F.** and **A. Torralba** (2009). New continuous-time multibit sigma-delta modulators with low sensitivity to clock jitter. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **56**(1), 74–83.

14. **Dorrer, L.**, **F. Kuttner**, **A. Santner**, **C. Kropf**, **T. Hartig**, **P. Torta**, and **P. Greco**, A 2.2 mW, continuous-time sigma-delta ADC for voice coding with 95dB dynamic range in a 65nm CMOS Process. *In Proceedings of the European Solid-State Circuits Conference*. 2006.

15. **Gao, W.**, **O. Shoaei**, and **W. Snelgrove**, Excess loop delay effects in continuous-time delta-sigma modulators and the compensation solution. *In Proceedings of IEEE International Symposium on Circuits and Systems*, volume 1. 1997.

16. **Gardner, F.** (1986). A transformation for digital simulation of analog filters. *IEEE Transactions on Communications*, **34**(7), 676–680.

17. **Gerfers, F.**, **M. Ortmanns**, and **Y. Manoli** (2003). A 1.5 V 12 bit power efficient continuous-time third order $\Sigma\Delta$ modulator. *IEEE Journal of Solid-State Circuits*, **38**(8), 1343–1352.

18. **Gray, R.** (1990). Quantization noise spectra. *IEEE Transactions on Information Theory*, **36**(6), 1220–1244.

19. **Hernandez, L.**, **A. Wiesbauer**, **S. Paton**, and **A. Di Giandomencio**, Modelling and optimization of low pass continuous-time sigma delta modulators for clock jitter noise reduction. *In Proceedings of IEEE International Symposium on Circuits and Systems*, volume 1. 2004.

20. **Inose, H.**, **Y. Yasuda**, and **J. Murakami** (1962). A telemetering system by code modulation. *IRE Transactions on Space Electronics and Telemetry*.

21. **Johns, D.** and **K. Martin**, *Analog integrated circuit design*. Wiley-India, 2009.

22. **Keller, M.**, **A. Buhmann**, **J. Sauerbrey**, **M. Ortmanns**, and **Y. Manoli** (2008). A comparative study on excess-loop-delay compensation techniques for continuous-time sigma–delta modulators. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **55**(11), 3480–3487.

23. **Kim, M.**, **G. Ahn**, **P. Hanumolu**, **S. Lee**, **S. Kim**, **S. You**, **J. Kim**, **G. Temes**, and **U. Moon** (2008). A 0.9 V 92 dB double-sampled switched-RC delta-sigma audio ADC. *IEEE Journal of Solid-State Circuits*, **43**(5), 1195–1206.

24. **Leuciuc, A.**, On the nonlinearity of integrators in continuous-time delta-sigma modulators. *In Proceedings of IEEE Midwest Symposium on Circuits and Systems*, volume 2. 2001.

25. **Loke, W.**, **C. Oh**, and **Y. Xu**, A 400-MHZ 4th-order CT bandpass sigma delta modulator with automatic frequency tuning. *In Proceedings of IEEE International Workshop on VLSI Design and Video Technology*. 2005.

26. **Mitteregger, G.**, **C. Ebner**, **S. Mechnig**, **T. Blon**, **C. Holuigue**, and **E. Romani** (2006). A 20-mW 640-MHz CMOS continuous-time $\Sigma\check{2}206$ ADC with 20-MHz signal bandwidth, 80-dB dynamic range and 12-bit ENOB. *IEEE Journal of Solid-State Circuits*, **41**(12), 2641–2649.

27. **Nguyen, K.**, **R. Adams**, **K. Sweetland**, and **H. Chen** (2005). A 106 dB SNR hybrid oversampling analog-to-digital converter for digital audio. *IEEE Journal of Solid-State Circuits*, **40**(12), 2408–2415.

28. **Norsworthy, S.**, **R. Schreier**, **G. Temes**, *et al.*, *Delta-Sigma data converters: theory, design, and simulation*. IEEE press New York, 1997.

29. **Oliaei, O.**, Clock jitter noise spectra in continuous-time delta-sigma modulators. *In Proceedings of IEEE International Symposium on Circuits and Systems*, volume 2. 1999.

30. **Oliaei, O.** (2003). Sigma-delta modulator with spectrally shaped feedback. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, **50**(9), 518–530.

31. **Ortmanns, M.**, **F. Gerfers**, and **Y. Manoli**, A continuous-time sigma-delta modulator with switched capacitor controlled current mode feedback. *In Proceedings of the European Solid-State Circuits Conference*. 2003.

32. **Ortmanns, M.**, **F. Gerfers**, and **Y. Manoli** (2005). A continuous-time sigma delta modulator with reduced sensitivity to clock jitter through SCR feedback. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **52**(5), 875–884.

33. **Ouzounov, S.**, **R. van Veldhoven**, **C. Bastiaansen**, **K. Vongehr**, **R. van Wegberg**, **G. Geelen**, **L. Breems**, and **A. van Roermund** (2007). A 1.2 V 121-mode CTDSM for wireless receivers in 90-nm CMOS. *Proceedings of the International Solid-State Circuits Conference*, 238–239.

34. **Papoulis, A.** and **S. Pillai**, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, 2002.

35. **Park, H.**, **K. Nam**, **D. Su**, **K. Vleugels**, and **B. Wooley** (2009). A 0.7 V 870 μW digital-audio CMOS sigma-delta modulator. *IEEE Journal of Solid-State Circuits*, **44**(4), 1078–1088.

36. **Pavan, S.** (2008). Excess loop delay compensation in continuous-time delta-sigma modulators. *IEEE Transactions on Circuits and Systems II: Express Briefs*, **55**(11), 1119–1123.

37. **Pavan, S.** (2010*a*). Efficient simulation of weak nonlinearities in continuous-time oversampling converters. *IEEE Transactions on Circuits and Systems I: Regular Papers*.

38. **Pavan, S.** (2010*b*). Systematic design centering of continuous time oversampling converters. *IEEE Transactions on Circuits and Systems II*, **57**, 158–162.

39. **Pavan, S.** and **N. Krishnapura** (2007). Automatic tuning of time constants in continuous-time delta–sigma modulators. *IEEE Transactions on Circuits and Systems II: Express Briefs*, **54**(4), 308–312.

40. **Pavan, S.**, **N. Krishnapura**, **R. Pandarinathan**, and **P. Sankar** (2008). A power optimized continuous-time $\Delta\Sigma$ converter for audio applications. *IEEE Journal of Solid-State Circuits*, **43**(2), 351–360.

41. **Pun, K.**, **S. Chatterjee**, and **P. Kinget** (2007). A 0.5 V 74 dB SNDR 25 kHz continuous-time delta-sigma modulator with a return-to-open DAC. *IEEE Journal of Solid State Circuits*, **42**(3), 496–507.

42. **Putter, B.**, $\Sigma\Delta$ ADC with finite impulse response feedback DAC. *In IEEE International Solid-State Circuits Conference*. 2004.

43. **Razavi, B.**, *Design of Analog CMOS Integrated Circuits*. Mcgraw-Hill Publication, 2005.

44. **Reddy, K.** and **S. Pavan** (2007). Fundamental limitations of continuous-time delta–sigma modulators due to clock jitter. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **54**(10), 2184–2194.

45. **Reddy, K.** and **S. Pavan**, A 20.7 mW continuous-time $\Delta\Sigma$ modulator with 15MHz bandwidth and 70 dB dynamic range. *In Proceedings of the European Solid-State Circuits Conference*. 2008.

46. **Roh, J.**, **S. Byun**, **Y. Choi**, **H. Roh**, **Y. Kim**, and **J. Kwon** (2008). A 0.9-V 60-μW 1-bit fourth-order delta-sigma modulator with 83-dB dynamic range. *IEEE Journal of Solid-State Circuits*, **43**(2), 361–370.

47. **Sankar, P.** and **S. Pavan** (2007). Analysis of integrator nonlinearity in a class of continuous-time delta–sigma modulators. *IEEE Transactions on Circuits and Systems II: Express Briefs*, **54**(12), 1125–1129.

48. **Saxena, S.**, **P. Sankar**, and **S. Pavan**, Automatic tuning of time constants in single bit continuous-time delta-sigma modulators. *In Proceedings of the International Symposium on Circuits and Systems*. 2009.

49. **Schreier, R.** and **G. Temes**, *Understanding delta-sigma data converters*. IEEE press Piscataway, NJ, 2005*a*.

50. **Schreier, R.** and **G. Temes**, *Understanding delta-sigma data converters*. IEEE Press - Piscataway, NJ, 2005*b*.

51. **Shoaei, O.** (1996). *Continuous-time delta-sigma A/D converters for high speed applications*. Ph.D. thesis, Carleton Unviersity.

52. **Silva, J.**, **U. Moon**, **J. Steensgaard**, and **G. Temes** (2001). Wideband low-distortion delta-sigma ADC topology. *Electronics Letters*, **37**, 737.

53. **Sripad, A.** and **D. Snyder** (1977). A necessary and sufficient condition for quantization errors to be uniform and white. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **25**(5), 442–448.

54. **Steensgaard, J.**, Nonlinearities in SC delta-sigma A/D converters. *In Proceedings of IEEE International Conference on Electronics, Circuits and Systems*, volume 1. 1998.

55. **Stevens, A.** and **G. Miller** (1994). A high-slew integrator for switched-capacitor circuits. *IEEE Journal of Solid-State Circuits*, **29**(9), 1146–1149.

56. **Tao, H.**, **L. Toth**, and **J. Khoury** (1999). Analysis of timing jitter in bandpass sigma-delta modulators. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, **46**(8), 991–1001.

57. **van der Zwan, E.** and **E. Dijkmans** (1996). A 0.2-mW CMOS $\Sigma\Delta$ modulator for speech coding with 80 dB dynamic range. *IEEE Journal of Solid-State Circuits*, **31**(12), 1873–1880.

58. **Widrow, B.** (1956). A study of rough amplitude quantization by means of Nyquist sampling theory. *IRE Transactions on Circuit Theory*, **3**(4), 266–276.

59. **Xia, B.**, **S. Yan**, and **E. Sanchez-Sinencio** (2004). An RC time constant auto-tuning structure for high linearity continuous-time $\Sigma\Delta$ modulators and active filters. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **51**(11), 2179–2188.

60. **Yan, S.** and **E. Sanchez-Sinencio** (2004). A continuous-time sigma-delta modulator with 88-dB dynamic range and 1.1-MHz signal bandwidth. *IEEE Journal of Solid-State Circuits*, **39**(1), 75–86.

61. **Yao, L.**, **M. Steyaert**, and **W. Sansen** (2004). A 1 V 140 μW 88 dB audio sigma-delta modulator in 90 nm CMOS. *IEEE Journal of Solid State Circuits*, **39**(11), 1809–1818.

LIST OF PAPERS BASED ON THIS THESIS

INTERNATIONAL JOURNALS

1. **S. Pavan and P. Sankar**, Power reduction in continuous-time delta-sigma modulators using the assisted opamp technique, *IEEE Journal of Solid-State Circuits*, 45(7), July 2010, 1365–1379.

2. **P. Sankar and S. Pavan**, Analysis of integrator nonlinearity in a class of continuous-time delta–sigma modulators, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 54(12), December 2007, 1125–1129.

INTERNATIONAL CONFERENCE

1. **S. Pavan and P. Sankar**, A 110 μW single bit audio continuous-time oversampled converter with 92.5 dB dynamic range, in *Proceedings of European Solid-State Circuits Conference*, September 2009, 287–290.

CURRICULUM VITAE

1. **NAME** : Prabu Sankar Thirugnanam

2. **DATE OF BIRTH** : 17^{th} August 1983.

3. **EDUCATIONAL QUALIFICATIONS:**

2001 Bachelor of Engineering (B.E.)

Institution	:	College of Engineering, Guindy.
University	:	Anna University.
Specialization	:	Electonics and Communications Engineering.

Doctor of Philosophy (Ph.D.)

Institution	:	Indian Institute of Technology, Madras.
Registration date	:	03-08-2005.

DOCTORAL COMMITTEE

CHAIRPERSON:	Head of the Department, Department of Electrical Engineering.
GUIDE:	Dr. Y. Shanthi Pavan, Professor, Department of Electrical Engineering.
MEMBERS:	Dr. Vinita Vasudevan, Professor, Department of Electrical Engineering.
	Dr. Nagendra Krishnapura, Asst. Professor, Department of Electrical Engineering.
	Dr. Shankar Balachandran, Asst. Professor, Department of Computer Science Engineering.
	Dr. Babu Viswanathan, Professor, Department of Mechanical Engineering.